SFB 649 Discussion Paper 2015-008

Nonparametric change-point analysis of volatility

Markus Bibinger*
Moritz Jirak*
Mathias Vetter**

*Humboldt-Universität zu Berlin, Germany
** Philipps-Universität Marburg, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

SFB 649 ECONOMIC RISK BERLIN

Nonparametric change-point analysis of volatility

Markus Bibinger[1,*,†], **Moritz Jirak**[2,*,†], **Mathias Vetter**[3,†]

[1] *Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany, bibinger@math.hu-berlin.de*

[2] *Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany, jirak@math.hu-berlin.de*

[3] *Fachbereich Mathematik und Informatik, Philipps-Universität Marburg, Hans-Meerwein-Straße 6, 35032 Marburg, Germany, vetterm@mathematik.uni-marburg.de*

Abstract: This work develops change-point methods for statistics of high-frequency data. The main interest is the volatility of an Itô semi-martingale, which is discretely observed over a fixed time horizon. We construct a minimax-optimal test to discriminate different smoothness classes of the underlying stochastic volatility process. In a high-frequency framework we prove weak convergence of the test statistic under the hypothesis to an extreme value distribution. As a key example, under extremely mild smoothness assumptions on the stochastic volatility we thereby derive a consistent test for volatility jumps. A simulation study demonstrates the practical value in finite-sample applications.

AMS 2000 subject classifications: Primary 62M10; secondary 62G10.
JEL classification: Primary C12; secondary C14.
Keywords and phrases: high-frequency data, nonparametric change-point test, minimax-optimal test, stochastic volatility, volatility jumps.

1. Introduction

Change-point theory classically focuses on detecting one or several structural breaks in the trend of time series. Statistical methods to infer change-points have a long and rich history, dating back to the pioneering work of Page (1955). Prominent approaches as e.g. by Hinkley (1971), Pettitt (1980), Andrews (1993) or Bai and Perron (1998), among many others, provide statistical tests for the hypothesis of no change-point against the alternative that changes occur. Moreover, they allow for localization of change-points (estimation) and confidence intervals. Change-point methods usually rely on maximum statistics and exploit limit theorems from extreme value theory; see Csörgő and Horváth (1997) for an overview. Less focus has been laid on discriminating jumps from continuous motion in a nonparametric framework. Important exceptions are Müller (1992), Müller and Stadtmüller (1999), Spokoiny (1998) and Wu and Zhao (2007) in the framework of nonparametric regression analysis. The latter serves as an important point of orientation for this work.

Statistics of high-frequency data is concerned with discretizations of continuous-time stochastic processes, most generally Itô semi-martingales. The continuous part of an Itô semi-martingale is of the form

$$X_t = X_0 + \int_0^t a_s\,ds + \int_0^t \sigma_s\,dW_s\,, \tag{1}$$

defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ with a standard $(\mathcal{F}_t)$-Brownian motion W and adapted drift and volatility processes a and σ. One key topic is statistical inference on the volatility

*Financial support from the Deutsche Forschungsgemeinschaft via SFB 649 *Ökonomisches Risiko* and FOR 1735 *Structural Inference in Statistics: Adaptation and Efficiency* is gratefully acknowledged.

†We thank Marc Hoffmann for helpful remarks on testing hypotheses of smoothness classes.

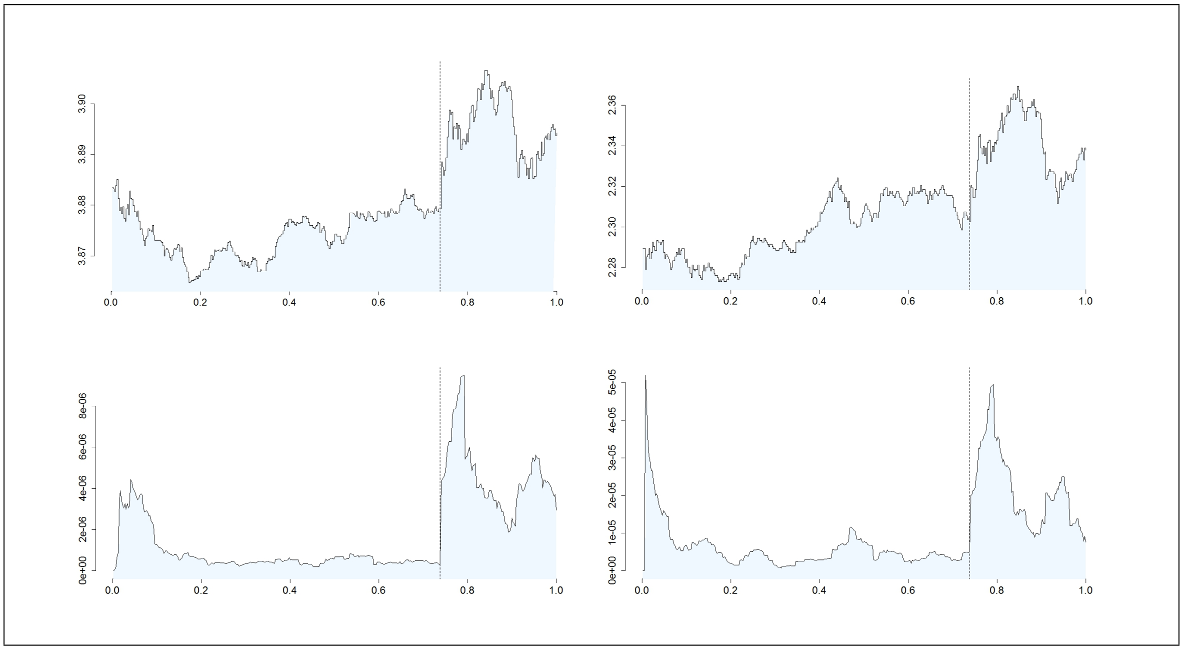

FIGURE 1. *Log-price intra-day evolutions (top) and estimated spot squared volatilities (bottom) for MMM (left) and GE (right) on March 18th, 2009.*

under high-frequency asymptotics when the mesh of a discretization on a fixed time horizon tends to zero. There is a vast body of works related to this problem and its economic implications; see e.g. Andersen and Bollerslev (1998), Mykland and Zhang (2009) and Jacod and Rosenbaum (2013), among many others. Statistics for a discretized continuous-time martingale is closely related to Gaussian calculus as highlighted by Mykland (2012) what is also at the heart of our analysis. Many contributions evolve around the question if jumps are present in the Itô semi-martingale modeling the log-price of a financial asset; see Aït-Sahalia and Jacod (2009) for a statistical test.

A more involved problem which is of key interest for economics and finance is to infer the smoothness of the underlying stochastic volatility process and to check whether volatility jumps occur. In particular, inference on volatility jumps allows to investigate the impact of certain news arrivals on financial risk. A first empirical study by Tauchen and Todorov (2011) indicates that volatility jumps can occur but, due to the lack of statistical methods, has been based on direct observations of the VIX, the most prominent available volatility index. Further contributions consider joint price-volatility jumps. Jacod and Todorov (2010) have designed a test to decide from high-frequency observations if contemporaneous jumps of an Itô semi-martingale and its volatility process have taken place at least once over some fixed time interval. These methods do not generalize to test directly for volatility jumps without restricting to a finite set of large price adjustments first. One main profit from our change-point analysis of high-frequency data is a general test for volatility jumps. Moreover, results on estimation of the time of a volatility jump are provided.

As an example, we illustrate in Figure 1 the evolution of log-prices of two blue-chip stocks, 3M and GE, over the NASDAQ intra-day trading period (6.5 h rescaled to the unit interval) on March 18th, 2009. We consider one minute returns from executed trades[1] to ensure the semi-martingale model is adequate and limit a manipulation by market microstructure frictions. Available tests and criteria

[1]reconstructed from the order book using LOBSTER, `https://lobster.wiwi.hu-berlin.de/`

do not identify price adjustments so large to be ascribed to jumps such that the test by Jacod and Todorov (2010) is not applicable. It seems as if a common source of news drives price dynamics at the end of that day concertedly. The picture becomes much clearer when focusing on the estimated spot squared volatilities in Figure 1, for which we average at each time point the previous 20 rescaled squared returns. This example suggests that volatility dynamics vary over time. Here, the volatilities of both assets sky-rocket at exactly the same time. This common volatility jump exactly coincides in time with a press release at 02:15 p.m. EST subsequent to a meeting of the Federal Open Market Committee. The time is marked in Figure 1 by the dashed lines. In light of increasing economic slack, the FOMC announced "to employ all available tools to promote economic recovery and to preserve price stability"[2], including a guarantee for an exceptionally low level of the federal funds rate for an extended period and a considerable increase of the size of the Federal Reserve's balance sheet. The mathematical concepts developed in this work provide a novel device for assessing volatility dynamics and jumps.

Quasi-likelihood estimation of a change-point in a diffusion parameter in a high-frequency setting has been considered by Iacus and Yoshida (2012), pointing out already one very useful bridge between change-point theory and high-frequency statistics. Our main focus is on testing for the presence of changes in a general setup exploiting localization techniques. Beyond the analysis of possible jumps of volatility there is great interest in the smoothness regularity of volatilities; see e.g. Gatheral et al. (2014) for a recent work, not least because of its crucial role for setting up volatility models.

We focus on volatilities which are almost surely locally bounded and strictly positive adapted processes. For our testing problem we consider classes of squared volatilities

$$\Sigma(\mathfrak{a}, L) = \Big\{ (\sigma_t^2(\omega))_{t\in[0,1]} \Big| \sup_{s,t\in[0,1],|s-t|<\delta} \big|\sigma_t^2(\omega) - \sigma_s^2(\omega)\big| \le L(\omega)\delta^{\mathfrak{a}} \Big\}\,, \tag{2}$$

for an almost surely bounded random variable L; cf. Assumption 3.1 for a precise statement about the conditions under the null hypothesis. The regularity exponent $\mathfrak{a} > 0$ is the key parameter to describe the null hypothesis H_0. We may now more formally ask the following questions:

(i) Is there a jump in the volatility, i.e. $\Delta\sigma_\theta^2 = \big(\sigma_\theta^2 - \lim_{s\uparrow\theta}\sigma_s^2\big) > 0$ for some $\theta \in (0,1)$?
(ii) Does volatility get rougher in the sense of a regularity exponent $\mathfrak{a}' < \mathfrak{a}$ on $(\theta, 1]$?
(iii) Can we discriminate different regularity exponents by a statistical test?

From a statistical perspective, the key question is which sizes of volatility jumps or which changes in the regularity exponent can be detected. For example, it is clear that we cannot detect jumps of arbitrarily small size. Loosely speaking, if we say that 'no jump' is our null hypothesis H_0, and 'there is a jump' our alternative H_1, then we face the problem of *distinguishability between* H_0 *and* H_1. The minimum size b_n of a jump $\Delta\sigma_\theta^2$, such that we are still able to uniformly control the type I and type II errors, is called *detection boundary*. If we are interested to test for the presence of jumps, we are thus led to consider for $\theta \in (0,1)$ alternatives of the form

$$\mathcal{S}_\theta^J(\mathfrak{a}, b_n, L) = \Big\{ (v_t)_{t\in[0,1]} \Big| (v_t - \Delta v_t)_{t\in[0,1]} \in \Sigma(\mathfrak{a}, L)\,;\, |\Delta v_\theta| \ge b_n \Big\} \tag{3}$$

with a decreasing sequence b_n. We then address the testing problem

$$H_0 : (\sigma_t^2(\omega))_{t\in[0,1]} \in \Sigma(\mathfrak{a}, L) \quad vs. \quad H_1 : \exists\,\theta \in (0,1) \text{ with } (\sigma_t^2(\omega))_{t\in[0,1]} \in \mathcal{S}_\theta^J(\mathfrak{a}, b_n, L)\,. \tag{4}$$

[2]source: www.federalreserve.gov/monetarypolicy/fomcminutes20090318.htm

In this context θ is commonly referred to as a *change-point*. The test alternative means that we demand at least one jump but do not exclude multiple jumps. The dependence on ω in (4) is natural in the definition of the hypotheses, as different realizations might lead to different paths on $[0, 1]$.

For the testing problem (4), we establish the minimax-optimal rate of convergence under high-frequency asymptotics. We follow the notion of minimax-optimality of statistical tests from the seminal contributions of Ingster (1993). For tests ψ that map a sample $\mathbf{X}_n$ to zero or one, where ψ accepts the null-hypothesis H_0 if $\psi = 0$ and rejects if $\psi = 1$, we consider the maximal type I error $\alpha_\psi\big(\mathfrak{a}\big) = \sup_{\sigma^2 \in \Sigma(\mathfrak{a}, L)} \mathbb{P}_\sigma\big(\psi = 1\big)$ and the maximal type II error $\beta_\psi\big(\mathfrak{a}, b_n\big) = \sup_{\sigma^2 \in \mathcal{S}_\theta^J(\mathfrak{a}, b_n, L)} \mathbb{P}_\sigma\big(\psi = 0\big)$ and define the global testing error as

$$\gamma_\psi\big(\mathfrak{a}, b_n\big) = \alpha_\psi\big(\mathfrak{a}\big) + \beta_\psi\big(\mathfrak{a}, b_n\big). \tag{5}$$

The primary interest now lies on tests that minimize $\gamma_\psi\big(\mathfrak{a}, b_n\big)$, given the boundary b_n. We aim to find sequences of tests ψ_n and boundaries b_n with the property that

$$\gamma_{\psi_n}\big(\mathfrak{a}, b_n\big) \to 0 \quad \text{as} \quad n \to \infty.$$

The smaller $b_n > 0$, the harder it is for a test to control the global testing error, i.e. to distinguish between H_0 and H_1. It is thus natural to pose the question given $\mathfrak{a}$, what is the minimal size of $b_n > 0$ such that

$$\lim_{n\to\infty} \inf_{\psi} \gamma_\psi\big(\mathfrak{a}, b_n\big) = 0 \tag{6}$$

holds? The optimal b_n^{opt} is called *minimax distinguishable boundary*, and a test ψ_n that satisfies (6) for all $b_n \geq b_n^{\text{opt}}$ *minimax-optimal*.

If L in (2) is deterministic, we prove that $b_n \propto (n/\log(n))^{\frac{-\mathfrak{a}}{2\mathfrak{a}+1}}$ constitutes the minimax distinguishable boundary for testing (4) and our constructed test is eligible to attain minimax-optimality. If $L(\omega)$ is only almost surely bounded, the rate sightly changes; see Section 4 for precise results. The particular form of this (lower) bound also appears in Spokoiny (1998), where the focus lies on nonparametric estimation of regression functions in the presence of jumps and where the jump size does not depend on n, see also Loader (1996). Yet, this estimation setup is very different from our situation.

For the lower bound proof we simplify the problem by information-theoretic reductions passing to more informative sub-classes of the parameter space. The lower bound established for the sub-class then serves a fortiori as a lower bound in the more general and less informative model. After gradually transforming the problem by showing strong Le Cam equivalences of the considered sub-experiment to more common situations with i.i.d. chi-square and Gaussian variables, the lower bound is proved by classical arguments based on the theory in Ingster and Suslina (2003).

Here we have outlined the theory with the focus on (4) and question **(i)**. Moreover, we shall consider minimax-optimal tests to address questions **(ii)** and **(iii)**. In the process, we may keep to a change-point setup with possible changes of the regularity exponent. However, the developed methods allow as well for tests of certain (global) smoothness regularities of the squared volatility.

The paper is organized as follows: Section 2 serves as an illustration for the benefit of cusum-based statistics in the simple, yet important model of a continuous Itô semimartingale with constant volatility. This illuminates the connection of classical change-point methods and high-frequency statistics. More involved, but also more important in practice is the case where the volatility is both time-varying and random. Section 3 is devoted to this nonparametric problem. As the volatility process is latent, which requires estimation based on smoothed squared increments of the semi-martingale, this poses an intricate statistical problem which to the best of the authors' knowledge had not been addressed so far. We establish a consistent test and derive a limit theorem under the hypothesis. The asymptotic analysis utilizes nonparametric change-point theory, stochastic calculus and bounds on the approximation error in the invariance principle. Our test allows to distinguish paths with jumps from continuous paths under remarkably general smoothness assumptions on the hypothesis. In Section 3.2 we discuss the situation in which the underlying Itô semimartingale might have jumps as well. Section 4 provides the theory on minimax-optimality with the lower bound, while Section 5 deals with the estimation of the location of the change in volatility under the alternative. A simulation study that investigates the finite-sample performance of the proposed methods and discusses some practical issues can be found in Section 6. All proofs are postponed to the Appendix.

2. Change-points in a parametric volatility model

Arguably, the simplest model of a continuous-time Itô diffusion X is the case of no drift and a constant volatility, such that X is given by

$$X_t = X_0 + \int_0^t \sigma \, dW_s \,, \tag{7}$$

where W denotes a standard Brownian motion. Throughout this work, the underlying process X is recorded at discrete regular times $i\Delta_n$ with a mesh $\Delta_n \to 0$. To keep the notation uncluttered, we assume to be on the fixed time interval $[0,1]$ and set $n = \Delta_n^{-1} \in \mathbb{N}$, so that we have observations $X_{i\Delta_n}, i = 0, \ldots, n$.

Inference on the squared volatility σ^2 is usually based on increments $\Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}$. In case one is interested in changes in the volatility, a natural quantity to discuss is the cusum statistic which reads

$$S_{n,m} = \frac{1}{\sqrt{n}} \sum_{i=1}^{m} \Big(n \big(\Delta_i^n X\big)^2 - \sum_{j=1}^{n} \big(\Delta_j^n X\big)^2 \Big) \,, m \in \{1, \ldots, n\} \,. \tag{8}$$

In order to derive the asymptotics of the cusum statistic, recall the functional (stable) central limit theorem for the realized volatility from observations of a continuous Itô semi-martingale (1) by Jacod (1997). Under mild assumptions, we have

$$\sqrt{n} \Bigg(\sum_{i=1}^{\lfloor nt \rfloor} \big(\Delta_i^n X\big)^2 - \int_0^t \sigma_s^2 \, ds \Bigg) \to \int_0^t \sqrt{2} \sigma_s^2 \, dB_s \;, t \in [0,1] \,, \tag{9}$$

as $n \to \infty$ weakly in the Skorokhod space with a standard Brownian motion B independent of W. In particular, if $\sigma_s = \sigma$ is constant, this result directly implies

$$S_{n,\lfloor nt \rfloor} \to \gamma \big(B_t - t B_1\big) \,, \tag{10}$$

with $\gamma^2 = \lim_{n\to\infty} n \, \mathbb{V}\mathrm{ar}\big(\sum_{i=1}^n \big(\Delta_i^n X\big)^2\big) = 2\sigma^4$, which coincides with a standard cusum limit theorem in the vein of Phillips (1987). The quarticity estimator by Barndorff-Nielsen and Shephard (2002), $\hat{\gamma}^2 = (2n/3) \sum_{i=1}^n \big(\Delta_i^n X\big)^4$, may be used to obtain a self-normalizing version:

$$\Bigg(\frac{2n}{3} \sum_{i=1}^{n} \big(\Delta_i^n X\big)^4 \Bigg)^{-1/2} S_{n,\lfloor nt \rfloor} \to B_t - t B_1 \,, \tag{11}$$

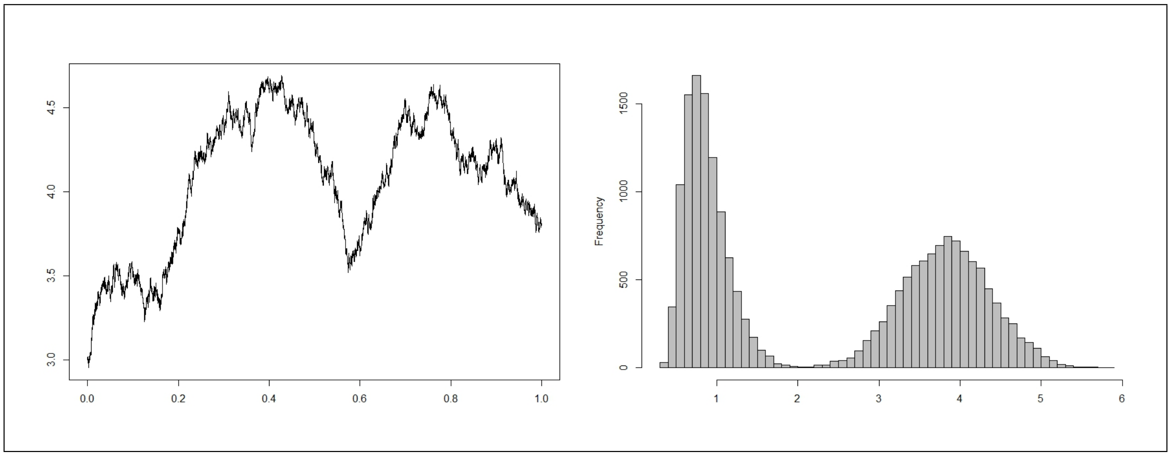

FIGURE 2. *Left: One realized path of X with structural break in volatility at $t = 1/2$. Right: Empirical results for the test statistic from 10000 iterations under both alternative and hypothesis.*

where the limit process is a standard Brownian bridge. Testing for jumps (resp. structural breaks, change-points) of the volatility is then pursued based on

$$T_n = \sup_{m=1,\ldots,n} \left| \left(\hat{\gamma}^2\right)^{-1/2} S_{n,m} \right| , \tag{12}$$

as the test statistic which (under the null that the volatility is constant) tends as $n \to \infty$ to a Kolmogorov-Smirnov law; see Marsaglia et al. (2003). Under the alternative T_n diverges almost surely.

Figure 2 shows an example in which we observe $n = 10000$ values of a standard Brownian motion under the hypothesis, while under the alternative the volatility jumps in $t = 1/2$ from 1 to 1.1. Out of 10000 Monte Carlo iterations for hypothesis and alternative, only 21 realizations of (12) under the hypothesis are larger than the minimum under the alternative. The other way round, in 11 iterations the values under the alternative fall below the maximum of the generated values from the hypothesis. The cusum approach hence clearly allows to separate hypothesis and alternative here, even for the relatively small volatility jump which is not readily identifiable from the path of X in Figure 2.

This is illustrated within the histogram in Figure 2. The left part stems from realizations under the hypothesis which closely track the asymptotic Kolmogorov-Smirnov law. The right part instead is due to realizations under the alternative. For larger volatility jumps the right part moves further to the right such that the two distributions separate even more clearly. This test based on (12) of Kolmogorov-Smirnov type permits to test the hypothesis of a constant volatility against structural breaks in an efficient way.

Beyond this bridging of classical change-point analysis and structural breaks in a parametric volatility model, our main focus in the sequel is nonparametric: to distinguish volatility jumps from a continuous motion of volatility or to identify changes in the regularity exponent.

3. Nonparametric change-point test for the volatility

3.1. Construction and limit behavior under the hypothesis

Suppose we observe a continuous Itô semi-martingale (1) at the regular times $i\Delta_n, i = 0, \ldots, n \in \mathbb{N}$. In this setting, we want to construct a test for (4). We write shortly

$$\big\{\omega|(\sigma_t^2(\omega))_{t\in[0,1]} \in \Sigma(\mathfrak{a}, L)\big\} = \Omega^c\,. \tag{13}$$

With the volatility process being time-varying, it becomes apparent from (9) that the test statistic (12) is not suitable to test H_0 against H_1. Our core idea is to utilize local two-sample t-tests over asymptotically small time blocks instead, thereby identifying breaks where segments are not smooth, in particular jumps, from too large deviations between two successive local estimators for the volatility.

As a first test statistic, we consider

$$V_n = \max_{i=0,\ldots,\lfloor n/k_n\rfloor-2} |X_{n,i}/X_{n,i+1} - 1|, \tag{14}$$

where $k_n \to \infty$ is an auxiliary sequence of integers depending on n and

$$X_{n,i} = \frac{n}{k_n}\sum_{j=1}^{k_n}(\Delta_{ik_n+j}^n X)^2\,, i = 0, \ldots, \lfloor n/k_n\rfloor - 1\,, \tag{15}$$

is a rescaled local version of realized volatility over blocks of the partition $[ik_n\Delta_n, (i+1)k_n\Delta_n]$. The $X_{n,i}$ estimate a block-wise constant proxy of the spot volatility $\sigma_{ik_n\Delta_n}$ on the respective blocks. Asymptotic properties of $X_{n,i}$ were e.g. derived in Alvarez et al. (2012). As mentioned above, a large distance between $X_{n,i}$ and $X_{n,i+1}$ suggests the presence of a jump or unsmooth breaks in the volatility close to time $ik_n\Delta_n$, which is why V_n appears to be a reasonable test statistic for our problem.

Our second test statistic is of the same nature as (14), but instead of non-overlapping blocks it takes into account all overlapping blocks of k_n increments:

$$V_n^* = \max_{i=k_n,\ldots,n-k_n}\left|\frac{\frac{n}{k_n}\sum_{j=i-k_n+1}^{i}(\Delta_j^n X)^2}{\frac{n}{k_n}\sum_{j=i+1}^{i+k_n}(\Delta_j^n X)^2} - 1\right|. \tag{16}$$

In comparison to nonparametric change-point approaches like the one by Wu and Zhao (2007), both statistics (14) and (16) are based on ratios rather than differences. This makes sense intuitively, since we are not dealing with the typical additive error structure of time series models. In our setting, we have e.g. $n(\Delta_i^n X)^2 \approx \sigma_{i\Delta_n}^2\chi_i^2, i = 1, \ldots, n$, with i.i.d. χ_1^2-distributed random variables χ_i^2, so that the volatility σ plays the role of a multiplicative error. Therefore, by computing ratios first, we basically deal with a maximum of identically distributed variables in the asymptotics. This is of key importance to obtain a distribution free limit under the hypothesis.

In order to discuss the asymptotics of V_n and V_n^* under the null-hypothesis we need a couple of additional assumptions, all of which are rather mild and are covered by a variety of stochastic volatility models.

Assumption 3.1. *The following assumptions on the processes a and σ are in order:*

(1) a and σ are locally bounded processes.
(2) σ is almost surely strictly positive, i.e. $\inf_{t\in[0,1]}\sigma_t^2 \geq \sigma_-^2 > 0$.

(3) On Ω^c, the modulus of continuity

$$w_\delta(\sigma)_t = \sup_{s,r\le t}\{|\sigma_s-\sigma_r| : |s-r|<\delta\}$$

is locally bounded in the sense that there exists $\mathfrak{a}>0$ and a sequence of stopping times $T_n\to\infty$ such that $w_\delta(\sigma)_{(T_n\wedge 1)}\le L_n\delta^{\mathfrak{a}}$, for some $\mathfrak{a}>0$ and some (a.s. finite) random variables L_n.

We choose the sequence $k_n\to\infty$, as $n\to\infty$, such that the following growth condition holds:

$$k_n^{-1}\Delta_n^{-\epsilon}+\sqrt{k_n}(k_n\Delta_n)^{\mathfrak{a}}\sqrt{\log(n)}\to 0\,, \tag{17}$$

for some $\epsilon>0$ and with $\mathfrak{a}>0$ from Assumption 3.1 (3).

There are two conditions contained in (17). First, $k_n\to\infty$ faster than some power of n which is a mild lower bound on the growth of k_n as $n\to\infty$. The second condition gives an upper bound related to the continuity of σ, since $\mathfrak{a}$ equals the Hölder index for an $\mathfrak{a}$-Hölder smooth volatility. Naturally, the smaller $\mathfrak{a}$ (and the less smooth σ), the smaller we have to choose the size of the blocks over which we estimate σ.

Theorem 3.2. *Set $m_n=\lfloor n/k_n\rfloor$ and $\gamma_{m_n}=[4\log(m_n)-2\log(\log(m_n))]^{1/2}$. If Assumption 3.1 holds and k_n satisfies condition (17), then we have on Ω^c (under H_0)*

$$\sqrt{\log(m_n)}\big(\big(k_n^{1/2}/\sqrt{2}\big)V_n-\gamma_{m_n}\big)\xrightarrow{w}V, \tag{18}$$

$$\sqrt{\log(m_n)}\big(k_n^{1/2}/\sqrt{2}\big)V_n^*-2\log(m_n)-\frac{1}{2}\log\log(m_n)-\log(3)\xrightarrow{w}V, \tag{19}$$

where V follows an extreme value distribution with distribution function

$$\mathbb{P}(V\le x)=\exp(-\pi^{-1/2}\exp(-x))\,. \tag{20}$$

Remark 3.3. It is remarkable that Theorem 3.2 in combination with condition (17) allows asymptotically to distinguish between volatility paths with jumps and volatility paths without jumps, where we only require some granted smoothness $\mathfrak{a}>0$ in Assumption 3.1 (3). Note that less smooth paths require smaller block lengths k_n by (17) which reduces the rate in Theorem 3.2 and the power of the test. Most importantly, we can cope with standard models for σ. For a continuous semi-martingale volatility, we have $\mathfrak{a}\approx 1/2$. In this case, we take $k_n\propto n^{1/2-\epsilon}$ for $\epsilon>0$ and ϵ small to preserve the highest possible power. Similarly, for a Lipschitz volatility, i.e. $\mathfrak{a}=1$, one might choose $k_n\propto n^{2/3-\epsilon}$. $\square$

As we show in Theorem 4.3 that V_n^* and V_n diverge under the alternative almost surely, Theorem 3.2 provides a consistent test with asymptotic power 1 by critical values from the limit law under the hypothesis.

3.2. A test in presence of jumps in the observed process

In order to provide a valid approach for various economic applications an important aim is to account for possible jumps in the process X as well. Thereto, consider a general Itô semi-martingale

$$X_t=X_0+\int_0^t a_s\,ds+\int_0^t\sigma_s\,dW_s+\int_0^t\int_{\mathbb{R}}\kappa(\delta(s,x))(\mu-\nu)(ds,dx) \tag{21}$$
$$+\int_0^t\int_{\mathbb{R}}\bar\kappa(\delta(s,x))\mu(ds,dx)\,,$$

where a truncation function κ, $\bar{\kappa}(x) = x - \kappa(x)$, separates large from compensated small jumps. The compensating intensity measure ν of the Poisson random measure μ admits the form $\nu(ds, dx) = ds \otimes \lambda(dx)$ for a σ-finite measure λ. Our notation follows Jacod (2008).

Assumption 3.4. *Grant Assumption 3.1 for the continuous part of X. Suppose $\sup_{\omega,x} |\delta(s,x)|/\gamma(x)$ is locally bounded for some deterministic non-negative function γ which satisfies for some $r < 2$:*

$$\int_{\mathbb{R}} (1 \wedge \gamma^r(x))\lambda(dx) < \infty\,. \tag{22}$$

In condition (22), r is a jump activity index that bounds the path-wise generalized Blumenthal-Getoor index from above. Imposing $r < 1$ restricts to jumps of finite variation and $r = 0$ to finite jump activity. To develop test statistics which are robust against jumps we employ a truncation principle as introduced for integrated volatility estimation by Mancini (2009) and Jacod (2008). The analogue of (14) with truncated squared increments reads

$$V_{n,u_n} = \max_{i=0,\dots,\lfloor n/k_n\rfloor-2} |X_{n,u_n,i}/X_{n,u_n,i+1} - 1|, \tag{23}$$

$$X_{n,u_n,i} = \frac{n}{k_n}\sum_{j=1}^{k_n} (\Delta^n_{ik_n+j}X)^2 \mathbb{1}_{\{|\Delta^n_{ik_n+j}X| \le u_n\}}\,, i = 0,\dots,\lfloor n/k_n\rfloor - 1\,. \tag{24}$$

The truncation sequence $u_n \propto n^{-\tau}$, $\tau \in (0, 1/2)$, is used to exclude large squared increments which can be ascribed to jumps. In the same way we can generalize statistic (16) with overlapping blocks:

$$V^*_{n,u_n} = \max_{i=k_n,\dots,n-k_n} \left| \frac{\frac{n}{k_n}\sum_{j=i-k_n+1}^{i} (\Delta^n_j X)^2 \mathbb{1}_{\{|\Delta^n_j X| \le u_n\}}}{\frac{n}{k_n}\sum_{j=i+1}^{i+k_n} (\Delta^n_j X)^2 \mathbb{1}_{\{|\Delta^n_j X| \le u_n\}}} - 1 \right|. \tag{25}$$

We prove below that truncation is an appropriate concept to asymptotically eliminate the influence by jumps, at least under certain restrictions on the jump activity, on k_n and on τ. In particular, under the hypothesis we obtain the same limit behaviour of the test statistics as in Theorem 3.2.

Proposition 3.5. *Suppose $k_n \propto n^\beta$ for $0 < \beta < 1$, such that condition (17) is satisfied. Furthermore, grant Assumption 3.4 for some*

$$r < \min\Big(2\big(2 - \tau^{-1}(1 - \beta/2)\big), \big(\tau^{-1}\min(1/2, 1 - \beta)\big), \big(2 - \tau^{-1}\beta/2\big)\Big) \tag{26}$$

as well. Then, with $m_n = \lfloor n/k_n\rfloor$ and $\gamma_{m_n} = [4\log(m_n) - 2\log(\log(m_n))]^{1/2}$ as before, and if either $r = 0$ or the jump process is a time-inhomogeneous Lévy process, we have on Ω^c (under H_0) for the statistics (23) and (25) the weak convergence

$$\sqrt{\log(m_n)}\Big(\big(k_n^{1/2}/\sqrt{2}\big)V_{n,u_n} - \gamma_{m_n}\Big) \xrightarrow{w} V, \tag{27}$$

$$\sqrt{\log(m_n)}\big(k_n^{1/2}/\sqrt{2}\big)V^*_{n,u_n} - 2\log(m_n) - \frac{1}{2}\log\log(m_n) - \log(3) \xrightarrow{w} V, \tag{28}$$

where V is distributed according to (20).

Remark 3.6. Arguably, it is most relevant from an applied perspective that the test based on (25) copes with finite activity jumps. In this case (26) reads as $\tau > 1/2 - \beta/4$, and the only requirement is that the threshold is not chosen too small. Beyond the finite activity case we restrict to time-inhomogenous Lévy jumps with independent increments where condition (26) becomes $r < 1$ for the typical case

$\beta \approx 1/2$ and $\tau \approx 1/2$. When $\tau \approx 1/2$ the third term in (26) is obsolete and the two others are equally weighted in case that $\beta \approx 1/2$. Thus, for other choices of β one makes (26) more restrictive, e.g. $r < 2/3$ for $\beta \approx 2/3$ under a Lipschitz volatility. The restrictions on the jump activity are stronger than what is usually needed for truncated realized volatility, i.e. $r < 1$ for general Itô semi-martingales; cf. Jacod (2008). This is due to the maximum in (25) compared to linear estimators. □

4. Consistency and minimax-optimal rate of convergence

In this section it becomes important that *stochastic* squared volatility processes lie under H_0 in $\Sigma(\mathfrak{a}, L_n)$, defined in (2), where we take into account strictly positive monotone increasing sequences L_n. This is crucial as we cannot describe the random processes as members of a fixed Hölder class. If σ_t^2 satisfies

$$\mathbb{E}\big[|\sigma_t^2 - \sigma_s^2|^{\mathfrak{b}}\big] \leq C\big|t-s\big|^{C+\mathfrak{b}\mathfrak{a}} \quad \text{for some } \mathfrak{b}, C > 1,$$

then the Kolmogorov-Čentsov Theorem implies that $\lim_{n\to\infty} \mathbb{P}\big((\sigma_t^2)_{0\leq t\leq 1} \in \Sigma(\mathfrak{a}, L_n)\big) = 1$, provided $L_n \to \infty$ arbitrarily slowly. Hence, up to a negligible set, $\Sigma(\mathfrak{a}, L_n)$ contains the *paths* generated by a huge number of popular volatility models when considering $L_n \to \infty$. On the other hand, if L is fixed, we are in the familiar framework of Hölder classes.

At this stage, we integrate alternatives where the volatility is less smooth than under the hypothesis, but which require not necessarily jumps. Note that the statistical devices developed above may be applied to discriminate H_0 from two kind of alternatives without jumps:

(R1) Until some change-point $\theta \in (0,1)$, the process $(\sigma_{t\wedge\theta}^2)$ behaves as a process in $\Sigma(\mathfrak{a}, L_n)$. After θ, the regularity exponent drops to some $0 < \mathfrak{a}' < \mathfrak{a}$. Since $\Sigma\big(\mathfrak{a}, L\big) \subset \Sigma\big(\mathfrak{a}', L\big)$, we require functions that 'exploit their roughness' in a certain sense; see below for a more detailed explanation. This change in the regularity exponent at time θ embeds into change-point theory. More precisely, we test the null of no change in regularity $\mathfrak{a}$ against a change at θ, where $\mathfrak{a}$ drops to $\mathfrak{a}'$.

(R2) Under H_0 the regularity exponent is $\mathfrak{a}$, and we would like to validate this hypothesis against the alternative hypothesis where the regularity is $\mathfrak{a}' < \mathfrak{a}$. Processes in the alternative set may be smoother on parts of the interval, but we require that they 'exploit their roughness' somewhere on $[0,1]$, such that in particular $(\sigma_t^2)_{t\in[0,1]} \notin \Sigma\big(\mathfrak{a}, L\big)$.

Though different, both problems are intimately connected. In fact, we demonstrate in the sequel that the same statistic can be applied to reach the optimal minimax rate. We also use the notation R1/2 to address both alternatives. To describe the alternative sets, define

$$\Delta_h^{\mathfrak{a}'} f_t = \frac{f_{t+h} - f_t}{|h|^{\mathfrak{a}'}}, \quad t \in [0,1], h \in [-1,1].$$

The terminology 'exploiting roughness' is specified and captured in the following sets. We then express the set of possible alternatives for *testing for a change in the regularity exponent*:

$$\mathcal{S}_\theta^{\mathrm{R1}}\big(\mathfrak{a}, \mathfrak{a}', b_n, L_n, C\big) = \Big\{\big(v_{t\wedge\theta}\big)_{t\in[0,1]} \in \Sigma(\mathfrak{a}, L_n)\big| \inf_{0\leq h\leq b_n^{1/\mathfrak{a}}} \Delta_h^{\mathfrak{a}'} v_\theta > C \text{ or } \sup_{0\leq h\leq b_n^{1/\mathfrak{a}}} \Delta_h^{\mathfrak{a}'} v_\theta < -C\Big\},$$

for some $C > 0$. If we want to *test for regularity* $\mathfrak{a}$ *in* σ_t^2 *against* $\mathfrak{a}' < \mathfrak{a}$, the set in question is slightly different:

$$\mathcal{S}_\theta^{\mathrm{R2}}\big(\mathfrak{a}, \mathfrak{a}', b_n, L_n, C\big) = \Big\{\big(v_t\big)_{t\in[0,1]} \in \Sigma(\mathfrak{a}', L_n)\big| \inf_{|h|\leq b_n^{1/\mathfrak{a}}} \Delta_h^{\mathfrak{a}'} v_\theta > C \text{ or } \sup_{|h|\leq b_n^{1/\mathfrak{a}}} \Delta_h^{\mathfrak{a}'} v_\theta < -C\Big\}.$$

In contrast to $\mathcal{S}_\theta^{\mathrm{R1}}$, volatilities in $\mathcal{S}_\theta^{\mathrm{R2}}$ need not to have regularity $\mathfrak{a}$ up to a change-point. On the other hand, it is crucial that we pose a condition on a centered window around θ here. We integrate both alternatives in the testing problem

$$H_0 : (\sigma_t^2(\omega))_{t\in[0,1]} \in \Sigma(\mathfrak{a}, L_n) \; vs. \; H_1^{\mathrm{R1/2}} : \exists\, \theta \in (0,1) \,|\, (\sigma_t^2(\omega))_{t\in[0,1]} \in \mathcal{S}_\theta^{\mathrm{R1/2}}\big(\mathfrak{a}, \mathfrak{a}', b_n, L_n, C\big). \quad (29)$$

Let us elaborate on the specific form of the alternative sets. In general, it is impossible to test $\Sigma(\mathfrak{a}, L_n)$ against $\Sigma(\mathfrak{a}', L_n)$ for $\mathfrak{a} > \mathfrak{a}'$, and it is necessary to consider special subsets of $\Sigma(\mathfrak{a}', L_n)$. Intuitively, it is clear that one needs at least to remove $\Sigma(\mathfrak{a}, L_n)$ from $\Sigma(\mathfrak{a}', L_n)$, but this is not sufficient. In fact, one needs to focus on the functions which exploit their roughness in the sense of the conditions in $\mathcal{S}_\theta^{\mathrm{R1/2}}$; cf. Hoffmann and Nickl (2011) for a detailed discussion in a related context. Geometrically, this means that the functions of interest are those which fluctuate considerably more, induced by demanding $\mathfrak{a}' < \mathfrak{a}$. However, as the sample size n grows, we only require this difference on sets that become smaller and smaller, which is expressed above in terms of b_n. Observe that for appropriate $C > 0$, for any $\theta \in (0,1)$ we have

$$\Big\{\mathcal{S}_\theta^{\mathrm{R1/2}}\big(\mathfrak{a}, \mathfrak{a}', b_n, L_n, C\big) \cap \Sigma(\mathfrak{a}', L_n)\Big\} \subseteq \Big\{(v_t)_{t\in[0,1]} \in \Sigma(\mathfrak{a}', L_n)\big| \inf_{\sigma^2\in\Sigma(\mathfrak{a},L_n)} \sup_{t\in[0,1]} |v_t - \sigma_t^2| \geq b_n\Big\},$$

where the right hand set has been used as alternative in Hoffmann and Nickl (2011). For the testing problem at hand, which is more complicated, this slightly larger set appears to be a too large alternative.

For the testing problems (4) and (29), we first present a negative result, that also serves as minimax lower bound for the problem depicted in (6).

Theorem 4.1. *Assume that $\mathfrak{a} > \mathfrak{a}' > 0$ and $\inf_t \sigma_t^2 \geq \sigma_-^2 > 0$. Consider either set of hypotheses $\{H_0, H_1^J\}$ or $\{H_0, H_1^{R1}\}$ or $\{H_0, H_1^{R2}\}$. Then for*

$$b_n \leq \big(n/\log(m_n)\big)^{-\frac{\mathfrak{a}}{2\mathfrak{a}+1}} \big(L_n \sigma_-^4\big)^{\frac{1}{2\mathfrak{a}+1}}\,, \quad (30)$$

we have in all three cases $\lim_{n\to\infty} \inf_\psi \gamma_\psi\big(\mathfrak{a}, b_n\big) = 1$.

Remark 4.2. Observe that for (29), the lower bound does not depend on the value of $\mathfrak{a}'$, only the fact that $\mathfrak{a}' < \mathfrak{a}$ is relevant. This is an asymptotic result though, and in practice the size of the difference $(\mathfrak{a} - \mathfrak{a}')$ may have a significant impact. □

Theorem 4.1 reveals that it is impossible to construct a minimax-optimal test in the sense of (6) if b_n is bounded as in (30). Consequently, we deduce that

$$b_n^{\mathrm{opt}} \geq \big(n/\log(m_n)\big)^{-\frac{\mathfrak{a}}{2\mathfrak{a}+1}} \big(L_n \sigma_-^4\big)^{\frac{1}{2\mathfrak{a}+1}}\,. \quad (31)$$

In Theorem 4.3 we shall establish a corresponding upper bound up to a multiplicative constant, and thus (31) already gives the optimal rate for the minimax distinguishable boundary. Observe that based on V_n^* from (16), we can obtain the following test $\psi^\diamond$.

$$\psi^\diamond\big((X_{i\Delta_n})_{0\le i\le n}\big) : \text{reject } H_0 \text{ if } V_n^* \geq 2C^\diamond \sqrt{2\log(m_n^\diamond)/k_n^\diamond}, \text{ i.e. } \psi^\diamond\big((X_{i\Delta_n})_{0\le i\le n}\big) = 1, \quad (32)$$

$$\text{where } C^\diamond > 2 \text{ and } \; k_n^\diamond = \Big(\sqrt{\log(m_n^\diamond)}\, n^{\mathfrak{a}}/L_n\Big)^{\frac{2}{2\mathfrak{a}+1}}, \quad m_n^\diamond = \lfloor n/k_n^\diamond \rfloor. \quad (33)$$

Alternatively, one might base a test on V_n from (14).

To simplify the discussion, we restrict to positive volatility jumps, i.e. $\inf_t \Delta\sigma_t \geq 0$, which appears natural from an economic point of view. We point out that an analogue result can be shown for negative, or positive and negative jumps, which however requires a further technical structural condition in case of multiple jumps in a vicinity for the alternative set.

Theorem 4.3. *Consider* (4) *with* $\inf_t \Delta\sigma_t \geq 0$*, or* (29) *with* $0 < \mathfrak{a}' < \mathfrak{a} \leq 1$ *and* $L_n = \mathcal{O}\big((n/k_n^\diamond)^{\mathfrak{a}-\mathfrak{a}'}\big)$*. If*

$$b_n^\diamond > \Big(2C^\diamond\sqrt{2\log(m_n^\diamond)/k_n^\diamond} + 2L_n\,(k_n^\diamond\Delta_n)^{\mathfrak{a}}\Big)\sup_{t\in[0,1]}\sigma_t^2\,, \tag{34}$$

where $k_n^\diamond$*,* $m_n^\diamond$ *and* $C^\diamond$ *are as in* (33)*, then* $\lim_{n\to\infty}\gamma_{\psi^\diamond}\big(\mathfrak{a}, b_n^\diamond\big) = 0$*. This implies that*

$$b_n^{opt} \propto \big(n/\log(n)\big)^{-\frac{\mathfrak{a}}{2\mathfrak{a}+1}} L_n^{1/(2\mathfrak{a}+1)}\,.$$

Remark 4.4. If L defined in (2) is deterministic, we get the minimax distinguishable boundary $b_n \propto (n/\log(n))^{\frac{-\mathfrak{a}}{2\mathfrak{a}+1}}$. □

5. Estimating the change-point

Once one has opted to reject the null hypothesis of no jump or break, the actual locations of jumps become of interest for further inference. This location problem has been extensively discussed in the literature in different frameworks; see for instance Csörgő and Horváth (1997) and Müller (1992). Here we develop a similar approach as in Aue et al. (2009). For the further analysis, we restrict ourselves to the 'one change-point alternative' involving a jump in the volatility, i.e. we specify the alternative hypothesis H_1^* as

$$H_1^*:\quad \big|\sigma_\theta - \sigma_{\theta-}\big| =: \delta \quad \text{for a unique } \theta\in(0,1).$$

To assess the possible time of change, we use slightly modified versions of the building blocks of the test statistic V_n^* from (16), defined as

$$V_{n,i}^\diamond = \frac{1}{\sqrt{k_n}}\bigg|\sum_{j=i-k_n+1}^{i} n(\Delta_j^n X)^2 - \sum_{j=i+1}^{i+k_n} n(\Delta_j^n X)^2\bigg|\,,$$

for $i = k_n,\ldots,n-k_n$, and $V_{n,i}^\diamond = 0$ else. In contrast to the construction of V_n^*, we may employ a simpler unweighted version. One can also consider the rescaled versions as in V_n^*, and the theoretical properties of these estimators coincide. The possible time of change is then estimated via

$$n\widehat{\theta}_n = \mathrm{argmax}_{i=k_n,\ldots,n-k_n} V_{n,i}^\diamond\,. \tag{35}$$

The following proposition establishes quantitative bounds for the quality of estimation.

Proposition 5.1. *Assume that the assumptions of Theorem 3.2 hold and that* H_1^* *is valid. Then, for* $\delta \geq 2k_n^{-1/2}\sqrt{\log(n)}\sup_{t\in[0,1]}\sigma_t^2$*, we have that*

$$\big|\widehat{\theta}_n - \theta\big| = \mathcal{O}_{\mathbb{P}}\bigg(\frac{\sqrt{k_n\log(n)}}{n\delta}\bigg)\,.$$

Remark 5.2. The estimator $\widehat{\theta}_n$ is consistent as long as $\sqrt{k_n \log(n)}/(n\delta) = \mathcal{O}(1)$. Note that jump sizes δ dependent on n may be considered, with $\delta = \delta_n \to 0$ as n increases. If δ does not tend to zero, the condition on δ in the proposition is always satisfied. The estimator extends to jumps of X using truncation as in (25), and Proposition 5.1 then applies to the generalized estimator under the assumptions of Proposition 3.5. □

Obviously, the quality of the estimator $\widehat{\theta}_n$ depends on the bandwidth k_n, and the smaller, the better. This is the complete opposite case compared to the test based on statistic V_n^*, where a larger choice of k_n increases the power. This is no contradiction, since both problems have a different, essentially reciprocal nature. Also note that k_n cannot be chosen arbitrarily small; see condition (17).
While classical estimators as the argmax of statistic (8) attain a standard $\sqrt{n}$-rate, corresponding to $k_n \approx n$, our nonparametric localization approach readily facilitates improved convergence rates as known for state-of-the-art change-point estimators. The following proposition sheds light on optimal convergence rates for the estimation problem.

Proposition 5.3. *On the assumptions of Proposition 5.1 for* $k_n \propto \big(\sqrt{\log(n)}n^{\mathfrak{a}}\big)^{\frac{2}{2\mathfrak{a}+1}}$, *a consistent estimator for* θ *does not exist in the case that* $\sqrt{k_n}\delta_n = \mathcal{O}\big(\sqrt{\log(n)}\big)$.

6. Simulations

We examine finite-sample properties of the proposed methods in a simulation study. First, consider $n = 10000$ observations at regular times of (1) with deterministic volatility function

$$\sigma_t = 1 - 0.01 \sin\left(\tfrac{3}{16}\pi\, t\right),\ t \in [0,1]\,, \tag{36}$$

with start value $X_0 = 4$ and with constant drift $a = 0.1$. We analyze the performance of the test with overlapping blocks based on test statistic V_n^* from (16) by simulating this model as hypothesis and add one jump of size 0.2 at fixed time $t = 0.425$ to σ_t as one specific alternative. Shifting the time of the volatility jump does not affect the results substantially. Also, alternatives with a rough but smooth movement of the same range result in similar effects.

The function (36) mimics a realistic volatility shape with strong decrease after opening and slight increase before closing and poses an intricate setup to discriminate jumps from continuous motion based on the $n = 10000$ discrete recordings of X. We have taken a jump size under the alternative which equals the range of the continuous movement. This means, that under the alternative the volatility jumps back at $t = 0.425$ to its maximum start value. This is in line with effects evoked by surprise elements from macroeconomic news in the financial context; see for instance Figure 1. Particularly, a decreasing continuous movement of volatility after opening (and slight U-shape) and a positive volatility jump appear as a realistic setup to us.

We have implemented the methods for various block sizes k_n, and all simulation results involve 5000 Monte Carlo iterations. We focus on V_n^* with overlapping blocks as it significantly outperforms the test with non-overlapping blocks using (14). For $k_{10000} = 500$, Figure 3 confirms a high finite-sample accuracy of the test. The empirical distribution under the hypothesis is remarkably close to the theoretical limit distribution. Besides, the power for an 0.05-level test is almost 95%. The configuration $k_{10000} = 500$ thus guarantees high power and a perfect fit by the limit theorem under the hypothesis.

Minor modifications of k_n do not change the results substantially. However, when we choose k_n much smaller, the power deteriorates. On the other hand, choosing k_n much larger increases the power close to 1 – but at the same time the fit by (19) becomes less accurate, as expected from condition (17).

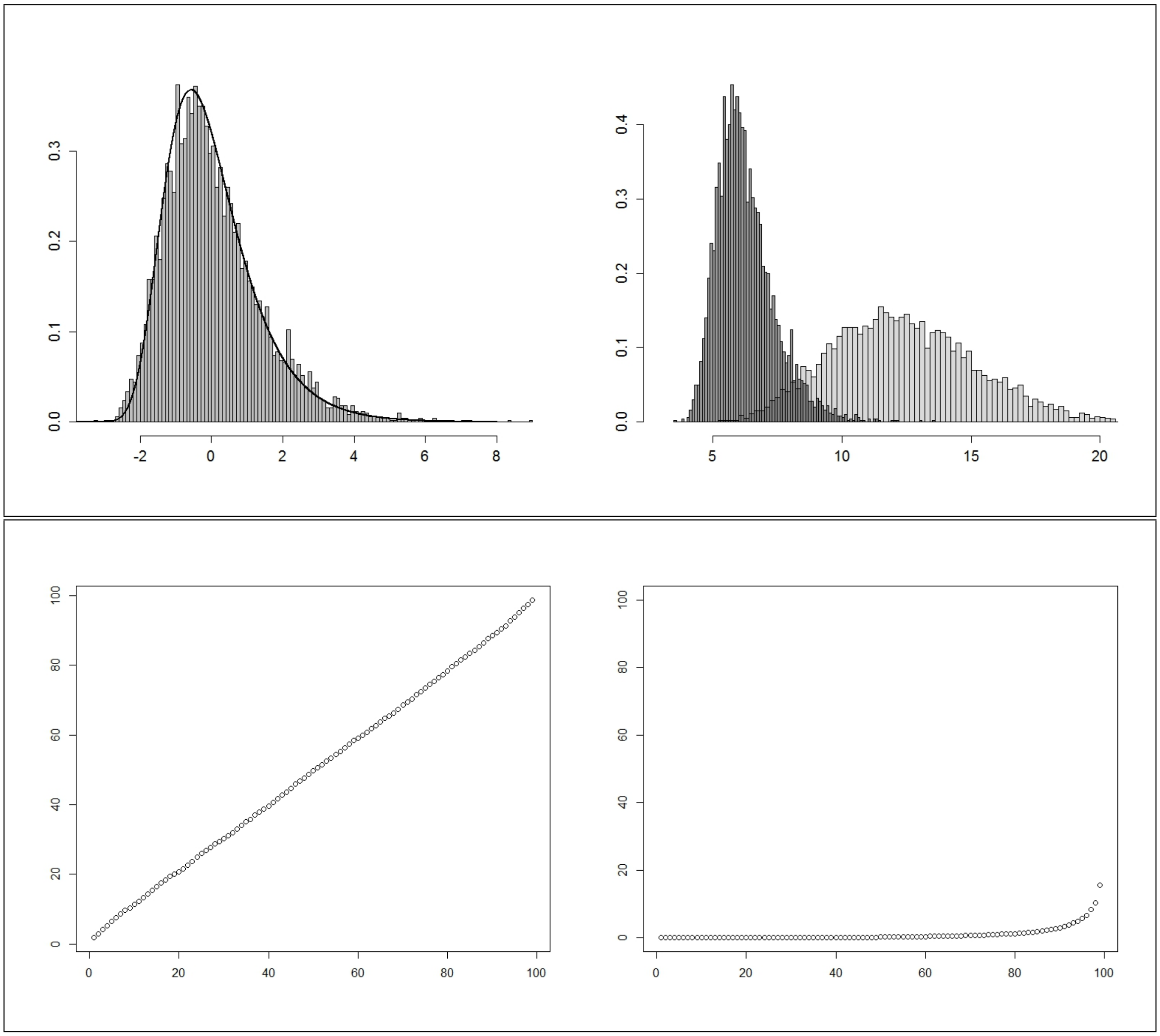

FIGURE 3. *Top: Histograms of* (16) *for* $k_{10000} = 500$ *under hypothesis and alternative (right) and rescaled version comparing left hand side and limit law of* (19) *(left); limit law density marked by solid line. Bottom: Empirical size (left) and power (right) of the test by comparing empirical percentiles to ones of limit law under* H_0.

This is illustrated in Figure 4 for $k_{10000} = 1000$. As limit theorems with extreme value distributions as limit laws are often imprecise in finite-sample applications, it is common practice in change-point literature to apply bootstrap-procedures; see e.g. Wu and Zhao (2007). This could be done here as well to access the law of (16) for larger k_n under the hypothesis when the fit by the limit law from (19) is not accurate. Such an approach allows to exploit that hypothesis and alternative separate even more clearly (we attain higher power) for such values of k_n.

In the sequel, we consider two further simulation experiments. In the first one, X additionally comprises jumps. Precisely, for the volatility as above in (36) one jump of X at a uniformly drawn jump arrival time is implemented for both the hypothesis and the alternative. Under the alternative X additionally exhibits a common jump of X and σ at $t = 0.425$. The jumps are $N(0.5, 0.1)$ distributed.

In the third simulation experiment, a stochastic volatility model is considered. We choose

$$v_t = \left(\int_0^t c \cdot \rho \, dW_s + \int_0^t \sqrt{1-\rho^2} \cdot c \, dW_s^\perp\right) \cdot \sigma_t \tag{37}$$

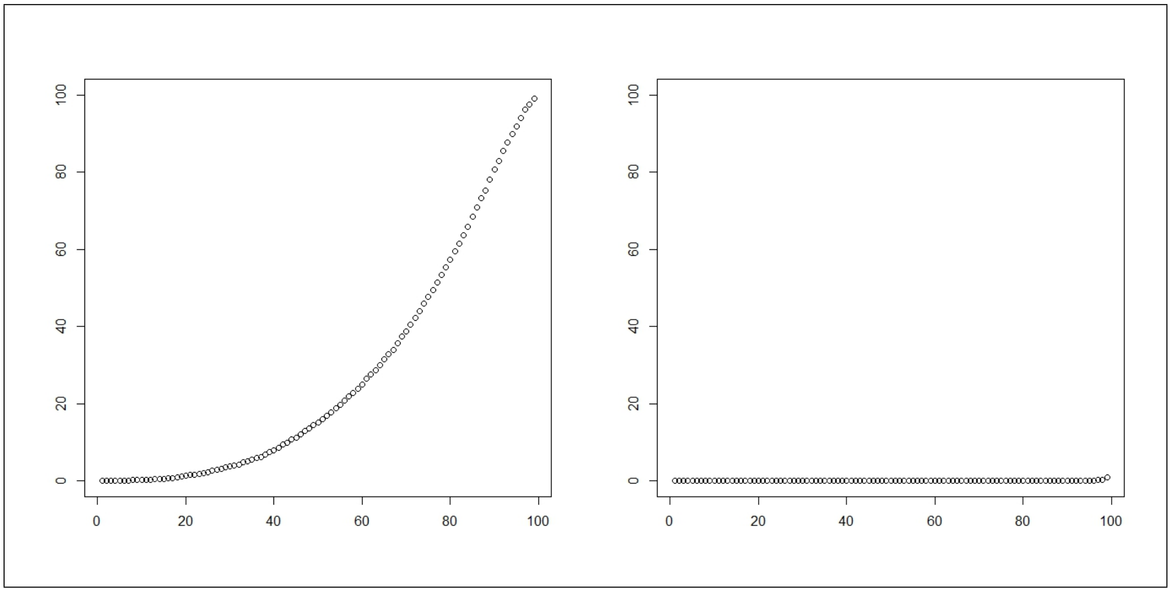

FIGURE 4. *Empirical size (left) and power (right) of the test by comparing empirical percentiles to ones of the limit law under* H_0, $k_{10000} = 1000$.

with $c = 0.1$ and $\rho = 0.5$, where $W^{\perp}$ is a standard Brownian motion independent of W and σ_t is the deterministic Lipschitz function (36) from above. Here, the volatility v_t takes the role of σ_t in (21) and follows a deterministic seasonality function, but for each path a random motion around that function is considered as well. The jumps are present in this experiment like in our second example.

In presence of jumps, X is of the form (21). Therefore, we apply the test statistic (25). For the truncation sequence we set $u_n = 2\log(n)n^{-1/2} \approx 18.42$ for $n = 10000$, motivated from extreme value theory and the magnitude of the increments from the continuous motion. In all cases, we simulate $n = 10000$ recordings and iterate 5000 Monte Carlo runs. Results for all three simulation scenarios are summarized in Table 1, i.e. the empirical power and size of tests with significance levels 1%, 5% and 10% are listed.

In all three experiments the power is very close to 1 for $k_n = 1000$ and slightly lower for $k_n = 500$. The fit of the empirical distribution by the limit V in (20) under the hypothesis is highly accurate for $k_n = 500$ in all setups and slightly deteriorates when $k_n = 1000$. Furthermore, the semi-martingale component of volatility in the third experiment slightly reduces the performance in terms of empirical power and size. Still, results are promising in view of the challenging model discriminating a volatility described by a continuous Itô semi-martingale from one with a jump. Comparing with Figures 3 and 4, we may conclude that there are basically no differences between the results for V_n^* from (16) applied to a continuous semi-martingale and for V_{n,u_n}^* from (25) applied to a process with jumps.

We have investigated in simulations also the influence of the rescaling factor in (16) by local volatility estimates. Thereto, we have compared the estimator to an oracle version where we rescale in each iteration with the generated volatility path. Surprisingly, the oracle version does not attain significant higher power than the original test. For instance, in the random volatility experiment, the power of the level 0.01 oracle test is 83.29% for $k_n = 500$ and 96.90% for $k_n = 1000$ and thus similar as listed for the test in Table 1. The statistics $\sup_x |F_n(x) - F(x)|$, with F the distribution function of V from (20) and F_n the empirical distribution of realizations under the hypothesis, take values 0.1453 and 0.3702 for $k_n = 500, 1000$, respectively, for the oracle test and 0.1527 and 0.5102 for the original test. This

Scenario	k_n	level 0.01	level 0.05	level 0.10
1 power	500	84.96	94.58	97.30
1 size	500	99.12	95.04	89.46
1 power	1000	99.08	99.96	99.98
1 size	1000	98.96	91.86	81.14
2 power	500	84.46	94.32	97.02
2 size	500	98.80	93.86	88.52
2 power	1000	99.10	99.88	100
2 size	1000	98.98	91.90	80.66
3 power	500	82.12	91.84	94.54
3 size	500	95.36	87.30	79.10
3 power	1000	97.08	99.18	99.64
3 size	1000	92.66	70.32	52.88

TABLE 1

Empirical size and power of tests in %, each time from 5000 Monte Carlo iterations. Theoretical asymptotic size is $(1-\alpha)100$ *for level* α *and theoretical asymptotic power 100.*

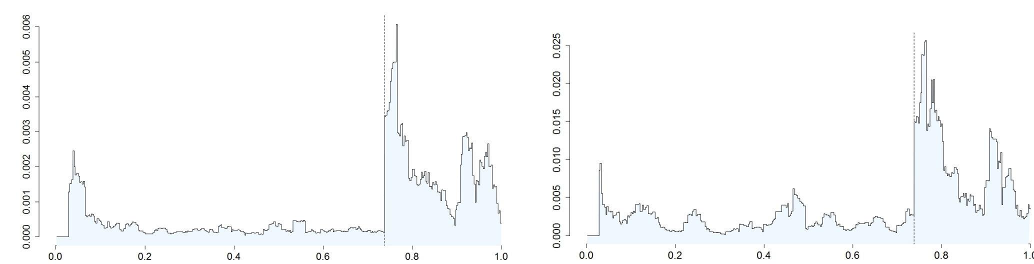

FIGURE 5. *Running local statistics for data example, MMM (left) and GE (right).*

indicates that the fit by the theoretical limit law is again just slightly more accurate for the oracle test.

Coming back to our introductory data example for intra-day prices on March 18th, 2009, we highlight in Figure 5 the evolution of the local rescaled averages over 10 squared returns, namely the numerator of the first term in the test statistic (16). The test rejects the null for both, 3M and GE with p-values very close to zero. The point in time where the difference of adjacent statistics is maximized estimates the timing of the structural change under the alternative. In both examples we find grid point 285 corresponding to 02:15 p.m. EST as estimated change-point.

Appendix A: Proof of Theorem 3.2

First, we reduce the proof of Theorem 3.2 to Propositions A.1-A.5. The main part in the analysis of V_n from (14) is to replace it by the statistic

$$U_n = \max_{i=0,\ldots,\lfloor n/k_n \rfloor -2} |Y_{n,i}/Y_{n,i+1} - 1|, \tag{38}$$

in which the original statistics $X_{n,i}$ from (15) are approximated by

$$Y_{n,i} = \frac{n}{k_n} \sum_{j=1}^{k_n} \sigma^2_{ik_n\Delta_n} (\Delta^n_{ik_n+j} W)^2. \tag{39}$$

Up to different (random) factors in front, the maximum in U_n is constructed from functionals of the i.i.d. increments of Brownian motion, which helps a lot in the derivation of its asymptotic behaviour. We start with a result on the approximation error due to replacing V_n by U_n.

Proposition A.1. *Suppose that we are under the null. If Assumption 3.1 and (17) hold, then we have*

$$\sqrt{\log{(n)}\, k_n}\, \big(V_n - U_n\big) \xrightarrow{\mathbb{P}} 0.$$

Recall that the variables $Y_{n,i}$ are not only computed over different intervals, but come with different volatilities in front as well. In order to obtain a statistic which is independent of σ let us define

$$\widetilde{Y}_{n,i} = \frac{n}{k_n} \sum_{j=1}^{k_n} \sigma^2_{(i-1)k_n\Delta_n} (\Delta^n_{ik_n+j} W)^2, \tag{40}$$

where the volatility factor is shifted in time now. Set then

$$\widetilde{U}_n = \max_{i=0,\ldots,\lfloor n/k_n \rfloor -2} |Y_{n,i}/\widetilde{Y}_{n,i+1} - 1|. \tag{41}$$

Proposition A.2. *Suppose that we are under the null. If Assumption 3.1 and (17) hold, then we have*

$$\sqrt{\log{(n)}\, k_n}\, \big(U_n - \widetilde{U}_n\big) \xrightarrow{\mathbb{P}} 0.$$

In the final step, we replace $\widetilde{Y}_{n,i+1}$ in the denominator by its limit $\sigma^2_{ik_n\Delta_n}$. Set

$$\widetilde{V}_n = \max_{i=0,\ldots,\lfloor n/k_n \rfloor -2} \Big|\frac{Y_{n,i} - \widetilde{Y}_{n,i+1}}{\sigma^2_{ik_n\Delta_n}}\Big|. \tag{42}$$

Proposition A.3. *Suppose that we are under the null. If condition (17) is satisfied, then we have*

$$\sqrt{\log{(n)}\, k_n}\, \big(\widetilde{U}_n - \widetilde{V}_n\big) \xrightarrow{\mathbb{P}} 0.$$

From Propositions A.1 to A.3 we have $\sqrt{\log{(n)}\, k_n}\, \big(V_n - \widetilde{V}_n\big) \xrightarrow{\mathbb{P}} 0$, while

$$\widetilde{V}_n = \max_{i=0,\ldots,\lfloor n/k_n \rfloor -2} \Big|\frac{1}{k_n}\sum_{j=1}^{k_n} (\sqrt{n}\Delta^n_{ik_n+j} W)^2 - \frac{1}{k_n}\sum_{j=1}^{k_n} (\sqrt{n}\Delta^n_{(i+1)k_n+j} W)^2\Big|. \tag{43}$$

This statistic corresponds to the statistic D_n given in (13) of Wu and Zhao (2007); see as well Proposition A.5. Precisely, after subtracting the mean on both sides above, their $(X_k)_{1\le k\le n}$ correspond to $\big((\sqrt{n}\Delta^n_k W)^2) - 1\big)_{1\le k\le n}$, which forms an i.i.d sequence of shifted χ^2_1-variables.

In the same fashion we can prove that the asymptotics of V_n^* in (16) can be traced back to the statistics D_n^* in (12) of Wu and Zhao (2007). See again also Proposition A.5.

Proposition A.4. *We have that* $\sqrt{\log(n)\,k_n}\,\big(V_n^* - \tilde{V}_n^*\big) \xrightarrow{\mathbb{P}} 0$*, with*

$$\widetilde{V}_n^* = \max_{i=k_n,\dots,n-k_n} \Big|\frac{1}{k_n}\sum_{j=i+1}^{i+k_n} \big((\sqrt{n}\Delta_j^n W)^2 - (\sqrt{n}\Delta_{j-k_n}^n W)^2\big)\Big|. \tag{44}$$

Theorem 1 of Wu and Zhao (2007) establishes limit theorems of the form (18) and (19) under more restrictive assertions on k_n than (17), as they consider the behavior for a class of weakly dependent random sequences $(X_k)_{k\geq 1}$. The next proposition provides a more specific limit theorem tailored to the asymptotic analysis of the statistics (43) and (44). In particular, instead of using the strong approximation theory under weak dependence from Wu (2007) employed by Wu and Zhao (2007) to prove their Theorem 1, we rely on classical bounds for the approximation error in the invariance principle for i.i.d. variables with existing moments. This is applicable in a more general setup with much smaller block lengths k_n.

Proposition A.5. *Consider a sequence* $(X_k)_{k\in\mathbb{N}}$ *of i.i.d. random variables with* $\mathrm{Var}\big[X_k\big] = \varsigma^2$ *and* $\mathbb{E}\big[|X_k|^p\big] < \infty$ *for some* $p \geq 4$*. If*

$$k_n^{-p/2} n = \mathcal{O}\big((\log(n))^{-p/2}\big), \tag{45}$$

then with $m_n = \lfloor n/k_n \rfloor$ *the statistic*

$$D_n^* = \frac{1}{k_n}\max_{k_n\leq i\leq n-k_n}\Big|\sum_{j=i+1}^{k_n+i} X_j - \sum_{j=i-k_n+1}^{i} X_j\Big|.$$

obeys the weak convergence:

$$\sqrt{\log(m_n)}(k_n^{1/2}\varsigma^{-1})D_n^* - 2\log(m_n) - \frac{1}{2}\log\log(m_n) - \log 3 \xrightarrow{w} V\,,$$

where V *is distributed according to* (20)*. The statistic*

$$D_n = \max_{1\leq i\leq \lfloor n/k_n\rfloor - 2}\big|\sum_{j=1}^{k_n} X_{ik_n+j} - X_{(i+1)k_n+j}\big|$$

using non-overlapping blocks satisfies under the same assumptions

$$\sqrt{\log(m_n)}\big((k_n^{1/2}\varsigma^{-1})D_n - [4\log(m_n) - 2\log(\log(m_n))]^{1/2}\big) \xrightarrow{w} V\,.$$

As all moments of the χ_1^2 distribution exist and k_n is at least of polynomial growth in n, Proposition A.5 applied to (43) and (44) implies Theorem 3.2. We start with the proof of Proposition A.5 and then show by proving Propositions A.1-A.4 that the preliminary reductions are in order.

Proof of Proposition A.5. By a simple rescaling argument, we can restrict ourselves to the case $\mathrm{Var}\big[X_k\big] = 1$. The Donsker-Prokhorov invariance principle guarantees weak convergence of partial sums of $(X_k)_{k\in\mathbb{N}}$, rescaled with $\sqrt{n}$, to the law of the standard Brownian motion as $n \to \infty$. Let $\big(Z_j\big)_{j\in\mathbb{N}}$ be a sequence of centered i.i.d. Gaussian random variables with $\mathbb{E}\big[Z_j^2\big] = \mathbb{E}\big[X_j^2\big] = 1$. Observe that

$$\max_{k_n\leq i\leq n-k_n}\Big|\sum_{j=i+1}^{k_n+i}(X_j - Z_j) - \sum_{j=i-k_n+1}^{i}(X_j - Z_j)\Big| \leq 4\max_{k_n\leq i\leq n}\Big|\sum_{j=1}^{i}(X_j - Z_j)\Big|.$$

We exploit the classical theory on bounds for the approximation error of partial sums of the above type associated with the invariance principle provided by the seminal works of Komlós et al. (1975), Komlós et al. (1976), Zaĭtsev (1987), and related literature. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence with $x_n \geq 0$ for all n. According to Theorem 4 of Komlós et al. (1976) or equivalently (1.6) of Sakhanenko (1996) with Markov inequality, the sequence $\big(Z_j\big)_{j\in\mathbb{N}}$ can be constructed in such a manner that

$$\mathbb{P}\bigg(\max_{k_n\le i\le n}\Big|\sum_{j=1}^{i}(X_j-Z_j)\Big|\geq x_n\bigg)\leq C_1\frac{1}{x_n^p}\sum_{j=1}^{n}\mathbb{E}\big[|X_j|^p\big]\leq C_2\frac{n}{x_n^p}\,,$$

with constants C_1, C_2 which may depend on p. Selecting $x_n = \sqrt{k_n}\delta_n$ with $\delta_n = (\log(n))^{-1/2}$, the conditions to apply Theorem 4 of Komlós et al. (1976) are in order and we get from condition (45) and the above that

$$\max_{k_n\le i\le n}\Big|\sum_{j=i+1}^{k_n+i}(X_j-Z_j)-\sum_{j=i-k_n+1}^{i}(X_j-Z_j)\Big| = \mathcal{O}_{\mathbb{P}}\big(\sqrt{k_n}(\log(n))^{-1/2}\big). \tag{46}$$

Denote with $\mathbb{B}(k)=\sum_{j=1}^{k}Z_j$ and define

$$H(u)=\big(\mathbb{1}(0\le u<1)-\mathbb{1}(-1<u<0)\big)/\sqrt{2}.$$

Then by (46), it follows that

$$\begin{aligned}\frac{\sqrt{k_n}D_n^*}{\sqrt{2}} &= \frac{1}{\sqrt{2k_n}}\max_{k_n\le i\le n-k_n}\big|\mathbb{B}(i+k_n)-2\mathbb{B}(i)+\mathbb{B}(i-k_n)\big|+\frac{\mathcal{O}_{\mathbb{P}}(1)}{\sqrt{\log(n)}}\\ &= \frac{1}{\sqrt{k_n}}\sup_{s\in[k_n,n-k_n]}\Big|\int_{\mathbb{R}}H\left(\frac{s-u}{k_n}\right)d\mathbb{B}(u)\Big|+\frac{\mathcal{O}(R_n)}{\sqrt{k_n}}+\frac{\mathcal{O}_{\mathbb{P}}(1)}{\sqrt{\log(n)}},\end{aligned}$$

where $R_n = \sup\big\{|\mathbb{B}(u)-\mathbb{B}(u')| \;:\; u,u'\in[0,n],\, |u-u'|\le 1\big\} = \mathcal{O}_{\mathbb{P}}\big(\sqrt{\log(n)}\big)$ by standard properties of Brownian motion. Then, since $(\log(n))^6 = \mathcal{O}\big(k_n\big)$ by condition (45), we may apply the limit theorem from Lemma 2 in Wu and Zhao (2007) with $\alpha=1$, $D_{H,1}=3$ and $b_n=m_n^{-1}$ for $m_n=\lfloor n/k_n\rfloor$; see Definition 1 and Lemma 2 of Wu and Zhao (2007).

In the same manner, we can use that

$$\sqrt{k_n}\Bigg(\sum_{j=i+1}^{k_n+i}X_j-\sum_{j=i-k_n+1}^{i}X_j\Bigg)=\sqrt{k_n}\Bigg(\sum_{j=i+1}^{k_n+i}Z_j-\sum_{j=i-k_n+1}^{i}Z_j\Bigg)+\mathcal{O}_{\mathbb{P}}\big(\sqrt{\log\,(n)}\big)\,,$$

for $\mathbb{E}[X_j^2]=1$ with $\big(Z_j\big)_{j\in\mathbb{N}}$ again a sequence of centered i.i.d. Gaussian random variables. Lemma 1 of Wu and Zhao (2007) then ensures the limit theorem for non-overlapping blocks. This completes the proof of Proposition A.5. □

Proof of Proposition A.1. First, a standard argument as e.g. laid out in Section 4.4.1 in Jacod and Protter (2012) allows us to strengthen Assumption 3.1 and to assume that all local conditions are in fact global. That is, we assume without loss of generality that $|a_s|\le K$, $0<\sigma_-^2<\sigma_s^2<K$ and $w_\delta(\sigma)_1\le K\delta^{\mathfrak{a}}$ for a generic constant K.

Let $(a_i)_{i=1,\dots,m}$ and $(b_i)_{i=1,\dots,m}$ be arbitrary reals. Obviously, for an arbitrary i we have

$$|a_i|\le|a_i-b_i|+|b_i|\le\max_{i=1,\dots,m}|a_i-b_i|+\max_{i=1,\dots,m}|b_i|.$$

Therefore the inequality holds with the left hand side replaced with $\max_{i=1,\ldots,m}|a_i|$ as well. Applied to V_n and U_n it is simple to deduce

$$|V_n - U_n| \le \max_{i=0,\ldots,\lfloor n/k_n\rfloor-2} |X_{n,i}/X_{n,i+1} - 1 - (Y_{n,i}/Y_{n,i+1} - 1)| \tag{47}$$
$$\le \max_{i=0,\ldots,\lfloor n/k_n\rfloor-2} \left| X_{n,i}\left(\frac{1}{X_{n,i+1}} - \frac{1}{Y_{n,i+1}}\right)\right| + \max_{i=0,\ldots,\lfloor n/k_n\rfloor-2}\left|\frac{X_{n,i}-Y_{n,i}}{Y_{n,i+1}}\right|.$$

Let us begin with the second term on the right hand side above. For all $\epsilon > 0$ and all constants $D > 0$, we have

$$\mathbb{P}\left(\max_{i=0,\ldots,\lfloor n/k_n\rfloor-2}\left|\frac{\sqrt{k_n\log(n)}(X_{n,i}-Y_{n,i})}{Y_{n,i+1}}\right| > \epsilon\right)$$
$$\le \mathbb{P}\left(\max_{i=0,\ldots,\lfloor n/k_n\rfloor-2}\sqrt{k_n\log(n)}|X_{n,i}-Y_{n,i}|\cdot \max_{i=0,\ldots,\lfloor n/k_n\rfloor-2} 1/|Y_{n,i+1}| > \epsilon\right)$$
$$\le \mathbb{P}\left(\max_{i=0,\ldots,\lfloor n/k_n\rfloor-2}\sqrt{k_n\log(n)}|X_{n,i}-Y_{n,i}| > \frac{\epsilon}{D}\right) + \mathbb{P}\left(\max_{i=0,\ldots,\lfloor n/k_n\rfloor-2} 1/|Y_{n,i+1}| > D\right). \tag{48}$$

To keep the notation readable, here and below we use standard probabilities and expectations without an extra indication that we are on the set Ω^c.

Since we have $\sigma_t^2 \ge \sigma_-^2 > 0$, we can use the same arguments as in the proof of equation (22) in Vetter (2012) to derive

$$\mathbb{P}\left(\max_{i=0,\ldots,\lfloor n/k_n\rfloor-2} 1/|Y_{n,i+1}| > D\right) = \mathbb{P}\left(\min_{i=0,\ldots,\lfloor n/k_n\rfloor-2}|Y_{n,i+1}| < D^{-1}\right) \to 0$$

with e.g. $D^{-1} = \sigma_-^2/2$. The intuition behind this result is that the probability of a mean of k_n i.i.d. variables with all moments deviating too much from its expectation becomes exponentially small in k_n. A similar argument will be given in (56) later. Also, here we need that k_n is of (at least) polynomial growth, which is included in (17).

On the other hand, using Itô formula we obtain

$$\sqrt{k_n\log(n)}(X_{n,i}-Y_{n,i}) = \frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\sum_{j=1}^{k_n}\left((\Delta_{ik_n+j}^n X)^2 - \sigma_{ik_n\Delta_n}^2(\Delta_{ik_n+j}^n W)^2\right) \tag{49}$$
$$= \frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\left(\sum_{j=1}^{k_n} 2\int_{(ik_n+j-1)\Delta_n}^{(ik_n+j)\Delta_n}(X_s - X_{(ik_n+j-1)\Delta_n})a_s\,ds + \sum_{j=1}^{k_n}\int_{(ik_n+j-1)\Delta_n}^{(ik_n+j)\Delta_n}(\sigma_s^2 - \sigma_{ik_n\Delta_n}^2)\,ds\right)$$
$$+ \frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\sum_{j=1}^{k_n} 2\int_{(ik_n+j-1)\Delta_n}^{(ik_n+j)\Delta_n}\left((X_s - X_{(ik_n+j-1)\Delta_n})\sigma_s - (W_s - W_{(ik_n+j-1)\Delta_n})\sigma_{ik_n\Delta_n}^2\right)dW_s\,.$$

Using this decomposition, we split the discussion of

$$\mathbb{P}\left(\max_{i=0,\ldots,\lfloor n/k_n\rfloor-2}\sqrt{k_n\log(n)}|X_{n,i}-Y_{n,i}| > \epsilon/D\right)$$

into three parts. For the first term, observe that

$$\mathbb{P}\Big(\max_{i=0,\dots,\lfloor n/k_n\rfloor-2}\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\Big|\sum_{j=1}^{k_n}2\int_{(ik_n+j-1)\Delta_n}^{(ik_n+j)\Delta_n}(X_s-X_{(ik_n+j-1)\Delta_n})a_s\,ds\Big|>\epsilon/(3D)\Big)$$
$$\le\sum_{i=0}^{\lfloor n/k_n\rfloor-2}\mathbb{P}\Big(\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\Big|\sum_{j=1}^{k_n}2\int_{(ik_n+j-1)\Delta_n}^{(ik_n+j)\Delta_n}(X_s-X_{(ik_n+j-1)\Delta_n})a_s ds\Big|>\epsilon/(3D)\Big)$$
$$\le(\epsilon/(3D))^{-r}\sum_{i=0}^{\lfloor n/k_n\rfloor-2}\mathbb{E}\Big[\Big|\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\sum_{j=1}^{k_n}2\int_{(ik_n+j-1)\Delta_n}^{(ik_n+j)\Delta_n}(X_s-X_{(ik_n+j-1)\Delta_n})a_s\,ds\Big|^r\Big],$$

for all integers r. Applying a standard bound based on Jensen's and Minkowski's inequalities yields

$$\mathbb{E}\Big[\Big|\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}2\int_{(ik_n+j-1)\Delta_n}^{(ik_n+j)\Delta_n}(X_s-X_{(ik_n+j-1)\Delta_n})a_s\,ds\Big|^r\Big]$$
$$\le K_r\bigg(\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\int_{(ik_n+j-1)\Delta_n}^{(ik_n+j)\Delta_n}\mathbb{E}[|X_s-X_{(ik_n+j-1)\Delta_n}|^r]^{1/r}ds\bigg)^r.$$

K_r here and below denotes a generic constant depending on r. Burkholder-Davis-Gundy inequality gives for any $s\in[(ik_n+j-1)\Delta_n,(ik_n+j)\Delta_n]$:

$$\mathbb{E}[|X_s-X_{(ik_n+j-1)\Delta_n}|^r]\le K_r n^{-r/2}\,,$$
$$\mathbb{E}\Big[\Big|\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}2\int_{(ik_n+j-1)\Delta_n}^{(ik_n+j)\Delta_n}(X_s-X_{(ik_n+j-1)\Delta_n})a_s\,ds\Big|^r\Big]\le K_r(nk_n)^{-r/2}\log^{r/2}(n)\,.$$

We conclude that

$$\mathbb{P}\Big(\max_{i=0,\dots,\lfloor n/k_n\rfloor-2}\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\Big|\sum_{j=1}^{k_n}2\int_{(ik_n+j-1)\Delta_n}^{(ik_n+j)\Delta_n}(X_s-X_{(ik_n+j-1)\Delta_n})a_s\,ds\Big|>\epsilon/(3D)\Big) \tag{50}$$
$$\le(\epsilon/(3D))^{-r}\lfloor n/k_n\rfloor K_r k_n^{r/2}n^{-r/2}\log^{r/2}(n)\to 0$$

for $r>2$ arbitrary. Regarding the second term in (49), on Ω^c we have

$$\max_{i=0,\dots,\lfloor n/k_n\rfloor-2}\Big|\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\sum_{j=1}^{k_n}\int_{(ik_n+j-1)\Delta_n}^{(ik_n+j)\Delta_n}(\sigma_s^2-\sigma_{ik_n\Delta_n}^2)\,ds\Big|$$
$$\le\max_{i=0,\dots,\lfloor n/k_n\rfloor-2}\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\sum_{j=1}^{k_n}\int_{(ik_n+j-1)\Delta_n}^{(ik_n+j)\Delta_n}|\sigma_s^2-\sigma_{ik_n\Delta_n}^2|ds$$
$$\le\sqrt{k_n}w_{k_n\Delta_n}(\sigma)_1\sqrt{\log(n)}\le K\sqrt{k_n}(k_n\Delta_n)^{\mathfrak{a}}\sqrt{\log(n)}\,,$$

which converges to zero by (17). Observe that addends above involve interlacing time intervals such that for σ a continuous Itô semi-martingale the bound above applies with $\mathfrak{a}=1/2$ and is sharp.

Finally, we have the further decomposition

$$(X_s-X_{(ik_n+j-1)\Delta_n})\sigma_s-(W_s-W_{(ik_n+j-1)\Delta_n})\sigma^2_{ik_n\Delta_n}=\sigma_s\int_{(ik_n+j-1)\Delta_n}^s a_u\,du \tag{51}$$
$$+(\sigma_s-\sigma_{ik_n\Delta_n})\int_{(ik_n+j-1)\Delta_n}^s\sigma_u\,dW_u+\sigma_{ik_n\Delta_n}\int_{(ik_n+j-1)\Delta_n}^s(\sigma_u-\sigma_{ik_n\Delta_n})\,dW_u\,.$$

We proceed in a similar way as above:

$$\mathbb{P}\Big(\max_{i=0,\ldots,\lfloor n/k_n\rfloor-2}\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\Big|\sum_{j=1}^{k_n}2\int_{(ik_n+j-1)\Delta_n}^{(ik_n+j)\Delta_n}\sigma_s\int_{(ik_n+j-1)\Delta_n}^{s}a_u du dW_s\Big|>\epsilon/(9D)\Big)$$
$$\le(\epsilon/(9D))^{-r}\sum_{i=0}^{\lfloor n/k_n\rfloor-2}\mathbb{E}\Big[\Big|\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\sum_{j=1}^{k_n}2\int_{(ik_n+j-1)\Delta_n}^{(ik_n+j)\Delta_n}\sigma_s\int_{(ik_n+j-1)\Delta_n}^{s}a_u\,du\,dW_s\Big|^r\Big]. \quad (52)$$

Precisely, let $r=2m$ and set

$$c_s=\sum_{j=1}^{k_n}\sigma_s\int_{(ik_n+j-1)\Delta_n}^{s}a_u\,du1_{[(ik_n+j-1)\Delta_n,(ik_n+j)\Delta_n)}(s)\,.$$

Then we have in a similar way as before

$$\begin{aligned}
&\mathbb{E}\Big[\Big|\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\sum_{j=1}^{k_n}2\int_{(ik_n+j-1)\Delta_n}^{(ik_n+j)\Delta_n}\sigma_s\int_{(ik_n+j-1)\Delta_n}^{s}a_u\,du\,dW_s\Big|^{2m}\Big]\\
&=2^{2m}\Big(\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\Big)^{2m}\mathbb{E}\Big[\Big|\int_{ik_n\Delta_n}^{(i+1)k_n\Delta_n}c_s\,dW_s\Big|^{2m}\Big]\\
&=2^{2m}\Big(\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\Big)^{2m}\mathbb{E}\Big[\Big(\int_{ik_n\Delta_n}^{(i+1)k_n\Delta_n}c_s^2ds\Big)^{m}\Big]\\
&\le K_m\Big(\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\Big)^{2m}\Big(\int_{ik_n\Delta_n}^{(i+1)k_n\Delta_n}\mathbb{E}[c_s^{2m}]^{1/m}ds\Big)^{m}.
\end{aligned}$$

With

$$\mathbb{E}[c_s^{2m}]=\sum_{j=1}^{k_n}\mathbb{E}\Big[\sigma_s^{2m}\Big(\int_{(ik_n+j-1)\Delta_n}^{s}a_udu\Big)^{2m}\Big]1_{[(ik_n+j-1)\Delta_n,(ik_n+j)\Delta_n)}(s)\le K_m\Delta_n^{2m}\,,$$

we obtain

$$\begin{aligned}
\Big(\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\Big)^{2m}\Big(\int_{ik_n\Delta_n}^{(i+1)k_n\Delta_n}\mathbb{E}[c_s^{2m}]^{1/m}ds\Big)^{m}&\le K_m\Big(\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\Big)^{2m}(k_n\Delta_n^3)^m\\
&=K_m\Delta_n^m\log^{2m}(n).
\end{aligned} \quad (53)$$

By choosing m large enough, the term in (52) converges to zero. Similarly,

$$\begin{aligned}
&\mathbb{E}\Big[\Big|\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\sum_{j=1}^{k_n}2\int_{(ik_n+j-1)\Delta_n}^{(ik_n+j)\Delta_n}(\sigma_s-\sigma_{ik_n\Delta_n})\int_{(ik_n+j-1)\Delta_n}^{s}\sigma_u\,dW_udW_s\Big|^{2m}\Big]\\
&\le K_m\Big(\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\Big)^{2m}\Big(\int_{ik_n\Delta_n}^{(i+1)k_n\Delta_n}\Big(\sum_{j=1}^{k_n}\mathbb{E}\Big[(\sigma_s-\sigma_{ik_n\Delta_n})^{2m}\\
&\qquad\times\Big(\int_{(ik_n+j-1)\Delta_n}^{s}\sigma_udW_u\Big)^{2m}1_{[(ik_n+j-1)\Delta_n,(ik_n+j)\Delta_n)}(s)\Big]\Big)^{1/m}ds\Big)^{m}
\end{aligned}$$

$$
\begin{aligned}
&\leq K_m\Big(\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\Big)^{2m}\Big(\int_{ik_n\Delta_n}^{(i+1)k_n\Delta_n}\Big(\sum_{j=1}^{k_n}\mathbb{E}\Big[w_{k_n\Delta_n}(\sigma)_1^{2m} \\
&\qquad\qquad \times\Big(\int_{(ik_n+j-1)\Delta_n}^{s}\sigma_u dW_u\Big)^{2m}1_{[(ik_n+j-1)\Delta_n,(ik_n+j)\Delta_n)}(s)\Big]\Big)^{1/m}ds\Big)^m \\
&\leq K_m\Big(\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\Big)^{2m}\Big(\int_{ik_n\Delta_n}^{(i+1)k_n\Delta_n}\big((k_n\Delta_n)^{2m\mathfrak{a}}\Delta_n^m\big)^{1/m}ds\Big)^m\leq K_m(k_n\Delta_n)^{2m\mathfrak{a}}.
\end{aligned}
$$

The same upper bound is obtained for the third term in (51). Again, choosing m large enough yields convergence to zero. Altogether, we conclude

$$
\mathbb{P}\left(\max_{i=0,\ldots,\lfloor n/k_n\rfloor-2}\sqrt{k_n\log(n)}|X_{n,i}-Y_{n,i}|>\epsilon/D\right)\to 0,
$$

and we are done with the second term on the right hand side of (47). Next, consider the first term on the right hand side of (47). For any $\epsilon>0$ and any $D>0$:

$$
\begin{aligned}
&\mathbb{P}\left(\max_{i=0,\ldots,\lfloor n/k_n\rfloor-2}\sqrt{k_n\log(n)}\big|X_{n,i}\left(\frac{1}{X_{n,i+1}}-\frac{1}{Y_{n,i+1}}\right)\big|>\epsilon\right) \\
&\leq \mathbb{P}\left(\max_{i=0,\ldots,\lfloor n/k_n\rfloor-2}\sqrt{k_n\log(n)}|X_{n,i}(Y_{n,i+1}-X_{n,i+1})|>\epsilon/D\right) \\
&\qquad\qquad +\mathbb{P}\left(\max_{i=0,\ldots,\lfloor n/k_n\rfloor-2}1/|Y_{n,i+1}X_{n,i+1}|>D\right).
\end{aligned}
$$

Observe that

$$
\begin{aligned}
&\mathbb{P}\left(\min_{i=0,\ldots,\lfloor n/k_n\rfloor-2}|Y_{n,i+1}X_{n,i+1}|<D^{-1}\right) \\
&\leq \mathbb{P}\left(\min_{i=0,\ldots,\lfloor n/k_n\rfloor-2}|Y_{n,i+1}|<D^{-1/2}\right)+\mathbb{P}\left(\min_{i=0,\ldots,\lfloor n/k_n\rfloor-2}|X_{n,i+1}|<D^{-1/2}\right) \\
&\leq \mathbb{P}\left(\min_{i=0,\ldots,\lfloor n/k_n\rfloor-2}|Y_{n,i+1}|<D^{-1/2}\right)+\mathbb{P}\left(\min_{i=0,\ldots,\lfloor n/k_n\rfloor-2}|Y_{n,i+1}|<2D^{-1/2}\right) \\
&\qquad\qquad +\mathbb{P}\left(\max_{i=0,\ldots,\lfloor n/k_n\rfloor-2}|X_{n,i+1}-Y_{n,i+1}|>D^{-1/2}\right).
\end{aligned}
$$

All three terms on the right hand side have already been discussed above for an appropriate choice of D, the latter term even with an additional factor $\sqrt{k_n\log(n)}$. Similarly, for all $\Gamma>0$ we have

$$
\begin{aligned}
&\mathbb{P}\left(\max_{i=0,\ldots,\lfloor n/k_n\rfloor-2}\sqrt{k_n\log(n)}|X_{n,i}(Y_{n,i+1}-X_{n,i+1})|>\epsilon/D\right) \\
&\leq \mathbb{P}\left(\max_{i=0,\ldots,\lfloor n/k_n\rfloor-2}|X_{n,i}|>\Gamma\right)+\mathbb{P}\left(\max_{i=0,\ldots,\lfloor n/k_n\rfloor-2}\sqrt{k_n\log(n)}|Y_{n,i+1}-X_{n,i+1}|>\epsilon/(D\Gamma)\right).
\end{aligned}
$$

Here we only have to focus on the first term, for which we use

$$
\begin{aligned}
&\mathbb{P}\left(\max_{i=0,\ldots,\lfloor n/k_n\rfloor-2}|X_{n,i}|>\Gamma\right) \\
&\leq \mathbb{P}\left(\max_{i=0,\ldots,\lfloor n/k_n\rfloor-2}|Y_{n,i}|>\Gamma/2\right)+\mathbb{P}\left(\max_{i=0,\ldots,\lfloor n/k_n\rfloor-2}|Y_{n,i+1}-X_{n,i+1}|>\Gamma/2\right).
\end{aligned}
\tag{54}
$$

The same arguments which were leading to equation (22) in Vetter (2012) show that the first probability becomes arbitrarily small for large enough Γ, this time because we may assume σ is bounded from above. The second probability has already been discussed above. $\square$

Proof of Proposition A.2. We have to show convergence in probability to zero of

$$\sqrt{k_n \log(n)} \max_{i=0,\dots,\lfloor n/k_n\rfloor-2} \Big| \frac{Y_{n,i}}{Y_{n,i+1}} - \frac{Y_{n,i}}{\widetilde{Y}_{n,i+1}} \Big| = \max_{i=0,\dots,\lfloor n/k_n\rfloor-2} \Big| \frac{\sqrt{k_n \log(n)} Y_{n,i}(\widetilde{Y}_{n,i+1} - Y_{n,i+1})}{Y_{n,i+1}\widetilde{Y}_{n,i+1}} \Big|.$$

Using equation (22) in Vetter (2012) again, we may focus on the numerator above only, and for the same reason as in (54) it suffices to prove convergence to zero of

$$\begin{aligned} &\mathbb{P}\left(\max_{i=0,\dots,\lfloor n/k_n\rfloor-2} \sqrt{k_n \log(n)} |\widetilde{Y}_{n,i+1} - Y_{n,i+1}| > \epsilon \right) \qquad (55) \\ &= \mathbb{P}\Big(\max_{i=0,\dots,\lfloor n/k_n\rfloor-2} \sqrt{k_n \log(n)} |\sigma^2_{ik_n\Delta_n} - \sigma^2_{(i+1)k_n\Delta_n}| \Big| \frac{n}{k_n} \sum_{j=1}^{k_n} (\Delta^n_{(i+1)k_n+j} W)^2 \Big| > \epsilon \Big) \\ &\le \mathbb{P}\left(\max_{i=0,\dots,\lfloor n/k_n\rfloor-2} \sqrt{k_n \log(n)} |\sigma^2_{ik_n\Delta_n} - \sigma^2_{(i+1)k_n\Delta_n}| > \epsilon/2 \right) \\ &\qquad\qquad + \mathbb{P}\left(\max_{i=0,\dots,\lfloor n/k_n\rfloor-2} \Big| \frac{n}{k_n} \sum_{j=1}^{k_n} (\Delta^n_{(i+1)k_n+j} W)^2 \Big| > 2 \right) \end{aligned}$$

for all $\epsilon > 0$. Regarding the first quantity, recall that on Ω^c by (17)

$$\begin{aligned} \max_{i=0,\dots,\lfloor n/k_n\rfloor-2} \sqrt{k_n \log(n)} |\sigma^2_{ik_n\Delta_n} - \sigma^2_{(i+1)k_n\Delta_n}| &\le \sqrt{k_n \log(n)} w_{k_n\Delta_n}(\sigma)_1 \\ &\le K\sqrt{k_n}(k_n\Delta_n)^{\mathfrak{a}} \sqrt{\log(n)} \to 0\,. \end{aligned}$$

On the other hand,

$$\begin{aligned} &\mathbb{P}\Big(\max_{i=0,\dots,\lfloor n/k_n\rfloor-2} \Big| \frac{n\sqrt{\log(n)}}{k_n} \sum_{j=1}^{k_n} (\Delta^n_{(i+1)k_n+j} W)^2 \Big| > 2 \Big) \qquad (56) \\ &\le \sum_{i=0}^{\lfloor n/k_n\rfloor-2} \mathbb{P}\Big(\Big| \frac{n\sqrt{\log(n)}}{k_n} \sum_{j=1}^{k_n} (\Delta^n_{(i+1)k_n+j} W)^2 \Big| > 2 \Big) \\ &\le \sum_{i=0}^{\lfloor n/k_n\rfloor-2} \mathbb{P}\Big(\Big| \frac{\sqrt{\log(n)}}{k_n} \sum_{j=1}^{k_n} \big((\sqrt{n}\Delta^n_{(i+1)k_n+j} W)^2 - 1 \big) \Big| > 1 \Big) \\ &\le \sum_{i=0}^{\lfloor n/k_n\rfloor-2} \mathbb{E}\Big[\Big| \frac{\sqrt{\log(n)}}{k_n} \sum_{j=1}^{k_n} \big((\sqrt{n}\Delta^n_{(i+1)k_n+j} W)^2 - 1 \big) \Big|^{2m} \Big] \end{aligned}$$

for all integers m. Due to the i.i.d. structure, the latter term is bounded by $K_m(n/k_n)k_n^{-m}\log^m(n)$, which converges to zero for m large enough. $\square$

Proof of Proposition A.3. We have to show convergence to zero in probability of

$$\sqrt{k_n\log(n)}\max_{i=0,\dots,\lfloor n/k_n\rfloor-2}\Big|\frac{Y_{n,i}-\widetilde{Y}_{n,i+1}}{\widetilde{Y}_{n,i+1}}-\frac{Y_{n,i}-\widetilde{Y}_{n,i+1}}{\sigma^2_{ik_n\Delta_n}}\Big|$$
$$=\sqrt{k_n\log(n)}\max_{i=0,\dots,\lfloor n/k_n\rfloor-2}\Big|\frac{(Y_{n,i}-\widetilde{Y}_{n,i+1})(\widetilde{Y}_{n,i+1}-\sigma^2_{ik_n\Delta_n})}{\widetilde{Y}_{n,i+1}\sigma^2_{ik_n\Delta_n}}\Big|.$$

It is sufficient to focus on the numerator, and we discuss two terms separately, using

$$\mathbb{P}\Big(\max_{i=0,\dots,\lfloor n/k_n\rfloor-2}\sqrt{k_n\log(n)}|(Y_{n,i}-\widetilde{Y}_{n,i+1})(\widetilde{Y}_{n,i+1}-\sigma^2_{ik_n\Delta_n})|>\epsilon\Big)$$
$$\le\mathbb{P}\Big(\max_{i=0,\dots,\lfloor n/k_n\rfloor-2}\sqrt{k_n\log(n)}|Y_{n,i}-\widetilde{Y}_{n,i+1}|>\sqrt{\epsilon}\Big)$$
$$+\,\mathbb{P}\Big(\max_{i=0,\dots,\lfloor n/k_n\rfloor-2}|\widetilde{Y}_{n,i+1}-\sigma^2_{ik_n\Delta_n}|>\sqrt{\epsilon}\Big)\,.$$

The first term has already been discussed in (55), while

$$\mathbb{P}\Big(\max_{i=0,\dots,\lfloor n/k_n\rfloor-2}|\widetilde{Y}_{n,i+1}-\sigma^2_{ik_n\Delta_n}|>\sqrt{\epsilon}\Big)$$
$$=\mathbb{P}\Bigg(\max_{i=0,\dots,\lfloor n/k_n\rfloor-2}\sigma^2_{ik_n\Delta_n}\Big|\frac{1}{k_n}\sum_{j=1}^{k_n}((\sqrt{n}\Delta^n_{(i+1)k_n+j}W)^2-1)\Big|>\sqrt{\epsilon}\Bigg)$$
$$\le\mathbb{P}\Bigg(\max_{i=0,\dots,\lfloor n/k_n\rfloor-2}\Big|\frac{1}{k_n}\sum_{j=1}^{k_n}((\sqrt{n}\Delta^n_{(i+1)k_n+j}W)^2-1)\Big|>\sqrt{\epsilon}/K\Bigg)$$

using $\sigma^2\le K$. The claim follows from (56). □

Proof of Proposition A.4. For the test statistic V_n^* from (16) our proof follows the same stages as the one for V_n via Propositions A.1, A.2 and A.3 above. We start proving

$$\sqrt{k_n\,\log(n)}\,(V_n^*-U_n^*)\xrightarrow{\mathbb{P}}0$$

for

$$U_n^*=\max_{i=k_n,\dots,n-k_n}\left|\frac{\frac{n}{k_n}\sum_{j=i-k_n+1}^{i}\sigma^2_{(i-k_n)\Delta_n}(\Delta_j^nW)^2}{\frac{n}{k_n}\sum_{j=i+1}^{i+k_n}\sigma^2_{i\Delta_n}(\Delta_j^nW)^2}-1\right|.$$

Similar to (47), we find that

$$|V_n^*-U_n^*|\le\max_{i=k_n,\dots,n-k_n}\Bigg|\sum_{j=i-k_n+1}^{i}(\Delta_j^nX)^2\Big(\Big(\sum_{j=i+1}^{i+k_n}(\Delta_j^nX)^2\Big)^{-1}-\Big(\sum_{j=i+1}^{i+k_n}\sigma^2_{i\Delta_n}(\Delta_j^nW)^2\Big)^{-1}\Big)\Bigg|$$
$$+\max_{i=k_n,\dots,n-k_n}\left|\frac{\frac{n}{k_n}\sum_{j=i-k_n+1}^{i}\big((\Delta_j^nX)^2-\sigma^2_{(i-k_n)\Delta_n}(\Delta_j^nW)^2\big)}{\frac{n}{k_n}\sum_{j=i+1}^{i+k_n}\sigma^2_{i\Delta_n}(\Delta_j^nW)^2}\right|.$$

Following an inequality analogous to (48), the key step is to show that

$$\mathbb{P}\Big(\sqrt{\log(n)k_n}\max_{i=k_n,\ldots,n-k_n}\Big|\frac{n}{k_n}\sum_{j=i-k_n+1}^{i}\big((\Delta_j^nX)^2-\sigma^2_{(i-k_n)\Delta_n}(\Delta_j^nW)^2\big)\Big|>\epsilon/D\Big)\to 0\,, \quad (57)$$

while for σ_t bounded from below we readily obtain

$$\mathbb{P}\Big(\min_{i=k_n,\ldots,n-k_n}\frac{n}{k_n}\Big|\sum_{j=i+1}^{i+k_n}\sigma^2_{i\Delta_n}(\Delta_j^nW)^2\Big|<D^{-1}\Big)\to 0\,.$$

For the proof of (57) we proceed with a decomposition analogous to (49) and for the first term along the same lines as above leading to (50). However, the maximum extends now over the larger set of all indices $i=k_n,\ldots,n-k_n$, and thus instead of (50) the upper bound yields $(\epsilon/(3D))^{-r}n^{1-\frac{r}{2}}k_n^{r/2}$ $\log^{r/2}(n)$, which is a factor k_n larger than above. Still, choosing r sufficiently large the term tends to zero. The same reasoning applies to all terms for which we have used Jensen's and generalized Minkowski's inequalities above as (53).

Upper bounds exploiting the smoothness of the volatility remain as before, for instance

$$\begin{aligned}\sqrt{\log(n)k_n}\max_{i=k_n,\ldots,n-k_n}\Big|\frac{n}{k_n}\sum_{j=1}^{k_n}\int_{(j+i-1)\Delta_n}^{(j+i+k_n-1)\Delta_n}(\sigma_s^2-\sigma^2_{i\Delta_n})\,ds\Big| &\le K\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\sum_{j=1}^{k_n}(j\Delta_n)^{\mathfrak{a}}\Delta_n\\ &\le K\sqrt{k_n}(k_n\Delta_n)^{\mathfrak{a}}\sqrt{\log(n)}\end{aligned}$$

with a constant K on Assumption 3.1. In this fashion, all terms generalizing the expressions in the proofs of Propositions A.1, A.2 and A.3 are controlled and we conclude Proposition A.4. □

Appendix B: Proof of Proposition 3.5

Recall the definition of the general Itô semi-martingale in (21). Again, by the usual localization procedure, we can work under the reinforced assumption that the process X_t and its jumps ΔX_t are bounded as well. We will then work with the decomposition $X_t=X_0+C_t+J_t$, where J_t denotes the pure jump martingale

$$J_t=\int_0^t\int_{\mathbb{R}}\delta(s,x)(\mu-\nu)(ds,dx)$$

and the continuous part becomes

$$C_t=\int_0^t\tilde a_s\,ds+\int_0^t\sigma_s\,dW_s$$

with $\tilde a_s=a_s+\int_{\mathbb{R}}\bar\kappa(\delta(s,x))\lambda(dx)$. The latter integral is finite for bounded jumps.

We shall prove only (28) of Proposition 3.5 by showing that

$$\begin{aligned}\sqrt{k_n\log(n)}\Bigg(\max_{i=k_n,\ldots,n-k_n}&\left|\frac{\frac{n}{k_n}\sum_{j=i-k_n+1}^{i}(\Delta_j^nX)^2\mathbb{1}_{\{|\Delta_j^nX|\le u_n\}}}{\frac{n}{k_n}\sum_{j=i+1}^{i+k_n}(\Delta_j^nX)^2\mathbb{1}_{\{|\Delta_j^nX|\le u_n\}}}-1\right|\\ &-\max_{i=k_n,\ldots,n-k_n}\left|\frac{\frac{n}{k_n}\sum_{j=i-k_n+1}^{i}(\Delta_j^nC)^2}{\frac{n}{k_n}\sum_{j=i+1}^{i+k_n}(\Delta_j^nC)^2}-1\right|\Bigg)\xrightarrow{\mathbb{P}}0\,. \quad (58)\end{aligned}$$

Following a decomposition of the error term as in (47), we have to show that

$$\sqrt{k_n\log(n)}\max_{i=k_n,\dots,n-k_n}\left|\frac{\frac{n}{k_n}\sum_{j=i-k_n+1}^{i}(\Delta_j^nX)^2\mathbb{1}_{\{|\Delta_j^nX|\le u_n\}}}{\frac{n}{k_n}\sum_{j=i+1}^{i+k_n}(\Delta_j^nX)^2\mathbb{1}_{\{|\Delta_j^nX|\le u_n\}}}-\frac{\frac{n}{k_n}\sum_{j=i-k_n+1}^{i}(\Delta_j^nX)^2\mathbb{1}_{\{|\Delta_j^nX|\le u_n\}}}{\frac{n}{k_n}\sum_{j=i+1}^{i+k_n}(\Delta_j^nC)^2}\right|$$
$$+\sqrt{k_n\log(n)}\max_{i=k_n,\dots,n-k_n}\left|\frac{\frac{n}{k_n}\sum_{j=i-k_n+1}^{i}\Big((\Delta_j^nX)^2\mathbb{1}_{\{|\Delta_j^nX|\le u_n\}}-(\Delta_j^nC)^2\Big)}{\frac{n}{k_n}\sum_{j=i+1}^{i+k_n}(\Delta_j^nC)^2}\right|\xrightarrow{\mathbb{P}}0\,.$$

Both terms are handled similarly and we restrict to the second one. It suffices to prove that

$$\mathbb{P}\bigg(\max_{i=k_n,\dots,n-k_n}\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}\Big|\sum_{j=i-k_n+1}^{i}\Big((\Delta_j^nX)^2\mathbb{1}_{\{|\Delta_j^nX|\le u_n\}}-(\Delta_j^nC)^2\Big)\Big|>\frac{\epsilon}{D}\bigg)\to0\,,\tag{59}$$

for all $\epsilon>0$ and constants $D>0$, as (58) then follows with (48) and the same bound for the second probability as in the proof of Proposition A.1. As $\max_{1\le i\le n}|\Delta_i^nC|=\mathcal{O}_{a.s.}(u_n)$ by basic extreme value theory we can work on a subset of Ω where $\max_{1\le i\le n}|\Delta_i^nC|=\mathcal{O}(u_n)$. Observe that on this subset

$$\max_{i=k_n,\dots,n-k_n}\Big|\sum_{j=i-k_n+1}^{i}(\Delta_j^nX)^2\mathbb{1}_{\{|\Delta_j^nX|\le u_n\}}-(\Delta_j^nC)^2\Big)\Big|$$
$$\le K\max_{i=k_n,\dots,n-k_n}\Big(\sum_{j=i-k_n+1}^{i}\mathbb{1}_{\{|\Delta_j^nX|>u_n\}}(\Delta_j^nC)^2+\sum_{j=i-k_n+1}^{i}\big((|\Delta_j^nJ|\wedge u_n)^2+(|\Delta_j^nJ|\wedge u_n)|\Delta_j^nC|\big)\Big)$$

with some constant K.

Pertaining the first addend and using $\max_i(\Delta_i^nC)^2=\mathcal{O}_{\mathbb{P}}(\Delta_n\log(n))$, we have to ensure that

$$\max_{i=k_n,\dots,n-k_n}\sum_{j=i-k_n+1}^{i}\mathbb{1}_{\{|\Delta_j^nX|>u_n\}}=\mathcal{O}_{\mathbb{P}}\big(\sqrt{k_n}/\log^{3/2}(n)\big)\,.$$

Let p with $1<p<(2r\tau)^{-1}$ be arbitrary. We use the decomposition $X=X'^n+X''^n$ with

$$X_t''^n=\int_0^t\int_{\mathbb{R}}\delta(s,x)\mathbb{1}_{\{\gamma(x)>u_n^p\}}\mu(ds,dx),\qquad X_t'^n=X_t-X_t''^n,$$

and define $A_j^n=\{|\Delta_j^nX'^n|\le u_n/2\}$. Finally, N^n is the counting process

$$N_t^n=\int_0^t\int_{\mathbb{R}}\mathbb{1}_{\{\gamma(x)>u_n^p\}}\mu(ds,dx).$$

We know from (13.1.10) in Jacod and Protter (2012) that

$$\mathbb{E}\Big[\max_{i=k_n,\dots,n-k_n}\sum_{j=i-k_n+1}^{i}\mathbb{1}_{\{|\Delta_j^nX|>u_n\}}\mathbb{1}_{\{(A_j^n)^\complement\}}\Big]\le\sum_{j=1}^{n}\mathbb{P}\big((A_j^n)^\complement\big)\to0$$

for all such p. Then, using $\mathbb{1}_{\{|\Delta_j^nX|>u_n\}}\mathbb{1}_{\{A_j^n\}}\le\mathbb{1}_{\{|\Delta_j^nX''^n|>u_n/2\}}$, all we have to show are conditions under which

$$\max_{i=k_n,\dots,n-k_n}\sum_{j=i-k_n+1}^{i}\mathbb{1}_{\{|\Delta_j^nN^n|\ge1\}}=\mathcal{O}_{\mathbb{P}}\big(\sqrt{k_n}/\log^{3/2}(n)\big)\,.$$

Obviously,

$$\sum_{j=i-k_n+1}^{i} \mathbb{1}_{\{|\Delta_j^n N^n|\ge 1\}} \le N^n_{i\Delta_n} - N^n_{(i-k_n)\Delta_n},$$

and N^n is a Poisson process with parameter $\int_{\mathbb{R}} \mathbb{1}_{\{\gamma(x)>u_n^p\}}\lambda(dx) \le K u_n^{-rp}$; see (13.1.14) in Jacod and Protter (2012). It is clearly enough if the probability of more than $l<\infty$ jumps on one block converges to zero, i.e.

$$\mathbb{P}\Big(\bigcup_{j=k_n}^{n}\big\{N^n_{(j+1)\Delta_n} - N^n_{(j-k_n+1)\Delta_n} \ge l\big\}\Big) \le n\mathbb{P}\big(N^n_{k_n\Delta_n}\ge l\big) \le Kn\Delta_n^l k_n^l u_n^{-rpl} = K k_n^l \Delta_n^{l(1-rp\tau)-1}\,.$$

Thus, we need the condition that for some $p>1$ and some $l<\infty$:

$$k_n^l\Delta_n^{l(1-rp\tau)-1}\to 0 \;\text{ and }\; 2r\tau<1\,. \tag{60}$$

Bounding the second term above comprising small jumps in case of non-truncation poses a more delicate mathematical problem. We restrict to the quadratic jump terms as the cross terms lead in the same way to an obsolete weaker criterion. For finite activity it is enough to ensure that $nk_n^{-1/2}\sqrt{\log(n)}u_n^2\to 0$. Else, define the sequence of random variables

$$\mathcal{Z}_i = \big(|\Delta_i J|\wedge u_n\big)^2 - \mathbb{E}\big[\big(|\Delta_i J|\wedge u_n\big)^2\big],\quad i=1,\ldots,n\,.$$

Note from equation (54) in Aït-Sahalia and Jacod (2010) that we can bound moments of $\big(|\Delta_i J|\wedge u_n\big)^2$ in the following way:

$$\begin{aligned}
&\mathbb{E}\Big[\big(|\Delta_i^n J|\wedge u_n\big)^2\big|\mathcal{F}_{(i-1)\Delta_n}\Big] \le K\Delta_n u_n^{2-r}\,,\\
&\mathbb{V}\mathrm{ar}\Big(\big(|\Delta_i^n J|\wedge u_n\big)^2\big|\mathcal{F}_{(i-1)\Delta_n}\Big) \le \mathbb{E}\Big[\big(|\Delta_i^n J|\wedge u_n\big)^4\big|\mathcal{F}_{(i-1)\Delta_n}\Big] \le u_n^2 K\Delta_n u_n^{2-r} = K\Delta_n u_n^{4-r}\,,
\end{aligned}$$

for all $i=1,\ldots,n$. We decompose

$$\begin{aligned}
\max_{i=k_n,\ldots,n-k_n}\Big|\sum_{j=i-k_n+1}^{i}\big(|\Delta_j J|\wedge u_n\big)^2\Big| \le \max_{i=k_n,\ldots,n-k_n}\Big|\sum_{j=i-k_n+1}^{i}\mathcal{Z}_j\Big|\\
+\max_{i=k_n,\ldots,n-k_n}\sum_{j=i-k_n+1}^{i}\mathbb{E}\Big[\big(|\Delta_j J|\wedge u_n\big)^2\Big]\,,
\end{aligned}$$

where the condition

$$\sqrt{k_n}u_n^{2-r}\sqrt{\log(n)}\to 0 \tag{61}$$

renders the second term with the expectation asymptotically negligible. Yet, the derivation of the maximum in the first term from its expectation can in general become much larger. Observe that

$$\max_{i=k_n,\ldots,n-k_n}\Big|\sum_{j=i-k_n+1}^{i}\mathcal{Z}_j\Big| = \max_{i=k_n,\ldots,n-k_n}\Big|\sum_{j=1}^{i}\mathcal{Z}_j - \sum_{j=1}^{i-k_n}\mathcal{Z}_j\Big| \le 2\max_{i=k_n,\ldots,n}\Big|\sum_{j=1}^{i}\mathcal{Z}_j\Big|\,.$$

Having a sequence of independent and centered random variables, we can apply Kolmogorov's maximal inequality:

$$\mathbb{P}\Big(\max_{i=k_n,\dots,n-k_n}\Big|\sum_{j=i-k_n+1}^{i}\mathcal{Z}_j\Big|>\lambda\Big)\le\frac{n}{\lambda^2}\mathbb{V}\mathrm{ar}(\mathcal{Z}_1)\le\lambda^{-2}u_n^{4-r}\,. \tag{62}$$

Thereby we conclude that $\max_{i=k_n,\dots,n-k_n}\Big|\sum_{j=i-k_n+1}^{i}\mathcal{Z}_j\Big|=\mathcal{O}_{\mathbb{P}}\big(u_n^{2-r/2}\big)$. We obtain the condition

$$\frac{n\sqrt{\log(n)}}{\sqrt{k_n}}u_n^{2-r/2}\to 0\,. \tag{63}$$

In conclusion, the conditions (60), (61) and (63) ensure (58), what yields our claim. □

Appendix C: Proof of the lower bound and consistency

Proof of Theorem 4.1. The proof is based on equivalences of statistical experiments in the strong Le Cam sense. After information-theoretic reductions, we subsequently move to statistical experiments that allow a simpler treatment; see (67) below. Our final experiment $\mathcal{E}_4$ is a special high-dimensional signal detection problem, from which we will deduce the lower bound by classical arguments.

First consider alternatives with a jump as in (4). Here, throughout this proof, we set

$$k_n=c_k\big(\sqrt{\log(m_n)}\sigma_-^2n^{\mathfrak{a}}/L_n\big)^{\frac{2}{2\mathfrak{a}+1}}\,, \tag{64}$$

with a constant $c_k>0$. In the preliminary step, we first grant the experimenter additional knowledge. We restrict to a sub-class of $\mathcal{S}_\theta^J\big(\mathfrak{a},b_n,L\big)$ from (3), where we have one jump at time $\theta\in(0,1)$ in the volatility, $|\sigma_\theta^2-\sigma_{\theta-}^2|\ge b_n$. Then, we assume that $\theta nk_n^{-1}\in\{1,2,\dots,\lfloor n/k_n\rfloor-1\}$, such that the jump time is in the set of observation grid points which are multiples of k_n. Furthermore, we can stick to $X_0=0$ and $a_s=0, s\in[0,1]$. From an information-theoretic view, obtaining this additional knowledge can only decrease the lower boundary on minimax distinguishability. Consequently, a lower bound derived for the sub-model carries over to the less informative general situation.

To ease the exposition, we first set $\sigma_-^2=1$ and $L_n=1$ and generalize the result at the end of this proof. Next, denote with $[a]_b=a \mod b$ and let

$$\sigma_{j\Delta_n}^2=\begin{cases}1+(k_n-[j]_{k_n})^{\mathfrak{a}}n^{-\mathfrak{a}}, & \theta n\le j<\theta n+k_n,\\ 1, & \text{else.}\end{cases} \tag{65}$$

The discretized squared volatility exhibits a jump (resp. change-point) of order b_n at θ and then decays on the window $[\theta,\theta+k_n\Delta_n]$ smoothly with regularity $\mathfrak{a}$ and is constant elsewhere. It suffices to consider the sub-class $\Sigma_\theta\subset\mathcal{S}_\theta^J\big(\mathfrak{a},b_n,L\big)$ of squared discretized volatility processes of the above form for which it remains unknown on which window the jump occurs.

Introduce a sequence r_n with $r_n\to\infty$ such that $r_nk_n^{-1}\to 0$ as $n\to\infty$. We specify the following stepwise approximation of $(\sigma_{j\Delta_n}^2)_{0\le j\le n}\in\Sigma_\theta$:

$$\widetilde{\sigma}_{j\Delta_n}^2=\begin{cases}1+(k_n-ir_n)^{\mathfrak{a}}n^{-\mathfrak{a}}, & \theta n+(i-1)r_n\le j\le\theta n+ir_n,\ 1\le i\le k_nr_n^{-1},\\ 1, & \text{else.}\end{cases} \tag{66}$$

Denote the observations by $\eta_j = \sigma_{(j-1)\Delta_n}\big(W_{j\Delta_n} - W_{(j-1)\Delta_n}\big)$ and $\widetilde{\eta}_j = \widetilde{\sigma}_{(j-1)\Delta_n}\big(W_{j\Delta_n} - W_{(j-1)\Delta_n}\big)$, $j = 1,\ldots,n$, respectively, with W the Wiener process in (1). In the sequel, it is convenient to distinguish the two cases where $\mathfrak{a} > 1/2$ and $\mathfrak{a} \le 1/2$.

Case $\mathfrak{a} > 1/2$: As alluded to above, we relate different experiments:

$\boldsymbol{\mathcal{E}_1}$: Observe $\big(\eta_j\big)_{1\le j\le n}$ and information $\theta n k_n^{-1} \in \{1,2,\ldots,\lfloor n/k_n\rfloor - 1\}$ is provided.

$\boldsymbol{\mathcal{E}_2}$: Observe $\big(\widetilde{\eta}_j\big)_{1\le j\le n}$ and information $\theta n k_n^{-1} \in \{1,2,\ldots,\lfloor n/k_n\rfloor - 1\}$ is provided.

$\boldsymbol{\mathcal{E}_3}$: Observe $\boldsymbol{\chi} = \big(\big(\widetilde{\sigma}^2_{ik_n\Delta_n}\chi_i\big)_{i\in\mathcal{I}_1}, \big(\widetilde{\sigma}^2_{\theta+(i-1)r_n\Delta_n}\tilde{\chi}_i\big)_{i\in\mathcal{I}_2}\big)$, where indices $(ik_n, i \in \mathcal{I}_1)$ expand over all multiples of k_n, except the one where the jump is located, i.e. $\mathcal{I}_1 = \{1,\ldots,\theta n k_n^{-1} - 1, \theta n k_n^{-1} + 1, \ldots, \lfloor n/k_n\rfloor - 1\}$, and $(\theta n + (i-1)r_n, i \in \mathcal{I}_2)$ over all multiples of r_n in the window of length $k_n\Delta_n$ where (σ_j^2) is non-constant, i.e. $\mathcal{I}_2 = \{1,2,\ldots,k_n r_n^{-1}\}$. $(\chi_i)_{i\in\mathcal{I}_1}$ and $(\tilde{\chi}_i)_{i\in\mathcal{I}_2}$ are i.i.d. random variables having chi-square distribution with degrees of freedom k_n for $i \in \mathcal{I}_1$ and r_n for $i \in \mathcal{I}_2$. Moreover, information $\theta n k_n^{-1} \in \{1,2,\ldots,\lfloor n/k_n\rfloor - 1\}$ is provided.

$\boldsymbol{\mathcal{E}_4}$: We observe $\boldsymbol{\xi} = \big(\big(k_n^{-1/2}\xi_i\widetilde{\sigma}^2_{ik_n\Delta_n} + \widetilde{\sigma}^2_{ik_n\Delta_n}\big)_{i\in\mathcal{I}_1}, \big(r_n^{-1/2}\tilde{\xi}_i\widetilde{\sigma}^2_{\theta+(i-1)r_n\Delta_n} + \widetilde{\sigma}^2_{\theta+(i-1)r_n\Delta_n}\big)_{i\in\mathcal{I}_2}\big)$ where $(\xi_i, \tilde{\xi}_i)$ are i.i.d. standard normal random variables. Moreover, information $\theta n k_n^{-1} \in \{1,2,\ldots,\lfloor n/k_n\rfloor - 1\}$ is provided.

When considering the above experiments, we always have $(\sigma^2_{j\Delta_n}) \in \varSigma_\theta$ (or $(\widetilde{\sigma}^2_{j\Delta_n}) \in \varSigma_\theta$) as unknown parameter that index a family of probability measures $\{\mathbb{P}_{(\sigma^2_{j\Delta_n})}\}$. For the sake of readability, we move this formalism to the background and omit subscripts indicating the parameter space. We show the following relations for the experiments, where $\sim$ marks strong Le Cam equivalence and $\approx$ asymptotic equivalence:

$$\mathcal{E}_1 \approx \mathcal{E}_2 \sim \mathcal{E}_3 \approx \mathcal{E}_4. \tag{67}$$

Finally, we shall derive the lower bound in $\boldsymbol{\mathcal{E}_4}$ which carries over to $\boldsymbol{\mathcal{E}_1}$ by the above relations and thus also to our general model. The proof is now divided into four main steps.

Step 1 $\mathcal{E}_1 \approx \mathcal{E}_2$: For random variables U, V and their laws $\mathbb{P}_U, \mathbb{P}_V$, we denote the Kullback-Leibler divergence $\mathbf{D}\big(U\|V\big) = \mathbf{D}\big(\mathbb{P}_U\|\mathbb{P}_V\big) = \int \log\big(d\mathbb{P}_U/d\mathbb{P}_V\big)d\mathbb{P}_U$. For normal families with unknown variance $\mathbb{P}_\theta = N(0,\theta)$, it is known that

$$\mathbf{D}\big(\mathbb{P}_\theta\|\mathbb{P}_{\theta'}\big) = \mathbb{E}_\theta\Big[\log\Big(\frac{d\mathbb{P}_\theta}{d\mathbb{P}_{\theta'}}\Big)\Big] = -\frac{1}{2}\Big(\log\Big(\frac{\theta}{\theta'}\Big) + 1 - \frac{\theta}{\theta'}\Big),$$

such that for $\theta = \theta' + \delta$ and considering asymptotics where $\delta \to 0$, we obtain

$$\mathbf{D}\big(\mathbb{P}_{\theta'+\delta}\|\mathbb{P}_{\theta'}\big) = -\frac{1}{2}\Big(\log\Big(1 + \frac{\delta}{\theta'}\Big) - \frac{\delta}{\theta'}\Big) = \frac{\delta^2}{4\theta'} + \mathcal{O}\big(\delta^3\big). \tag{68}$$

As $\mathcal{E}_1$ and $\mathcal{E}_2$ share a common space on which the considered random variables are accommodated, asymptotic equivalence holds if $\|\mathbb{P}_{(\eta_j)} - \mathbb{P}_{(\tilde{\eta}_j)}\|_{TV} \to 0$ as $n \to \infty$ where $\|\cdot\|_{TV}$ denotes the total variation distance and $\mathbb{P}_{(\eta_j)}$ the law of observations (η_j). We exploit Pinsker's inequality

$$\big\|\mathbb{P}_{(\eta_j)} - \mathbb{P}_{(\tilde{\eta}_j)}\big\|^2_{TV} \le \frac{1}{2}\mathbf{D}\big((\eta_j)\|(\tilde{\eta}_j)\big). \tag{69}$$

By Gaussianity and independence of Brownian increments, implying additivity of the Kullback-Leibler divergences, it follows with (68) for any piecewise constant approximation of a function with regularity $\mathfrak{a}$ on $k_n r_n^{-1}$ intervals of length $r_n\Delta_n$ that

$$\mathbf{D}\big((\eta_j)\|(\tilde{\eta}_j)\big)=\mathcal{O}(1)\sum_{i=1}^{k_n r_n^{-1}}\sum_{j=1}^{r_n}\big(j\Delta_n\big)^{2\mathfrak{a}}=\mathcal{O}\big(n^{-2\mathfrak{a}}k_n r_n^{2\mathfrak{a}}\big)\,,$$

which tends to zero for $r_n k_n^{-1}=\mathcal{O}(n^{-\epsilon})$ for some $\epsilon>0$.

Step 2 $\mathcal{E}_2\sim\mathcal{E}_3$: The vector of averages

$$\bigg(\Big(k_n^{-1}\sum_{j=1}^{k_n}\eta^2_{ik_n+j-1}\Big)_{i\in\mathcal{I}_1},\ \Big(r_n^{-1}\sum_{j=1}^{r_n}\eta^2_{\theta n+(i-1)r_n+j-1}\Big)_{i\in\mathcal{I}_2}\bigg)$$

forms a sufficient statistic for $(\widetilde{\sigma}^2_{j-1})_{1\le j\le n}$. Thereby we conclude, see e.g. Lemma 3.2 of Brown and Low (1996), the strong Le Cam equivalence.

Step 3 $\mathcal{E}_3\approx\mathcal{E}_4$: Let $\boldsymbol{\chi}^\diamond=\Big(k_n^{-1/2}\big(\widetilde{\sigma}^2_{ik_n\Delta_n}(\chi_i-k_n)\big)_{i\in\mathcal{I}_1},r_n^{-1/2}\big(\widetilde{\sigma}^2_{\theta+(i-1)r_n\Delta_n}(\tilde{\chi}_i-r_n)\big)_{i\in\mathcal{I}_2}\Big)$ and $\boldsymbol{\xi}^\diamond=\Big(\big(\xi_i\widetilde{\sigma}^2_{ik_n\Delta_n}\big)_{i\in\mathcal{I}_1},\big(\tilde{\xi}_i\widetilde{\sigma}^2_{\theta+(i-1)r_n\Delta_n}\big)_{i\in\mathcal{I}_2}\Big)$. In both experiments random variables are accommodated on the same space. Rescaling and a location shift yield with Pinsker's inequality

$$\big\|\mathbb{P}_{\boldsymbol{\chi}}-\mathbb{P}_{\boldsymbol{\xi}}\big\|_{TV}^2=\big\|\mathbb{P}_{\boldsymbol{\chi}^\diamond}-\mathbb{P}_{\boldsymbol{\xi}^\diamond}\big\|_{TV}^2\le\frac{1}{2}\mathbf{D}\big(\boldsymbol{\chi}^\diamond\|\boldsymbol{\xi}^\diamond\big)\,.$$

By independence, it follows that

$$\mathbf{D}\big(\boldsymbol{\chi}^\diamond\|\boldsymbol{\xi}^\diamond\big)\le\sum_{i\in\mathcal{I}_1}\mathbf{D}\Big(k_n^{-1/2}\widetilde{\sigma}^2_{\frac{ik_n}{n}}(\chi_i-k_n)\big\|\xi_i\widetilde{\sigma}^2_{\frac{ik_n}{n}}\Big)+\sum_{i\in\mathcal{I}_2}\mathbf{D}\Big(r_n^{-1/2}\widetilde{\sigma}^2_{\theta+\frac{(i-1)r_n}{n}}(\tilde{\chi}_i-r_n)\big\|\tilde{\xi}_i\widetilde{\sigma}^2_{\theta+\frac{(i-1)r_n}{n}}\Big).$$

An application of Theorem 1.1 in Bobkov et al. (2013) yields

$$\begin{aligned}&\sum_{i\in\mathcal{I}_1}\mathbf{D}\big(k_n^{-1/2}\widetilde{\sigma}^2_{ik_n\Delta_n}(\chi_i-k_n)\big\|\xi_i\widetilde{\sigma}^2_{ik_n\Delta_n}\big)=\mathcal{O}\big(nk_n^{-2}\big)\,,\\&\sum_{i\in\mathcal{I}_2}\mathbf{D}\big(r_n^{-1/2}\widetilde{\sigma}^2_{\theta+(i-1)r_n\Delta_n}(\tilde{\chi}_i-r_n)\big\|\tilde{\xi}_i\widetilde{\sigma}^2_{\theta+(i-1)r_n\Delta_n}\big)=\mathcal{O}\big(k_n r_n^{-2}\big)\,.\end{aligned}$$

For $\mathfrak{a}>1/2$, we have $nk_n^{-2}=\mathcal{O}(1)$. Choosing r_n sufficiently large such that $k_n r_n^{-2}=\mathcal{O}(1)$, it follows that

$$\big\|\mathbb{P}_{\boldsymbol{\chi}}-\mathbb{P}_{\boldsymbol{\xi}}\big\|_{TV}=\mathcal{O}(1)\,,\tag{70}$$

what ensures the claimed asymptotic equivalence.

Step 4: By the previous steps, it suffices to establish a lower bound for the distinguishability in experiment $\mathcal{E}_4$. Adding an additional drift, which gives clearly an equivalent experiment, we consider observations $\boldsymbol{\xi}=\big(\big(k_n^{-1/2}\xi_i\widetilde{\sigma}^2_{ik_n\Delta_n}+\widetilde{\sigma}^2_{ik_n\Delta_n}-1\big)_{i\in\mathcal{I}_1},\big(r_n^{-1/2}\tilde{\xi}_i\widetilde{\sigma}^2_{\theta+(i-1)r_n\Delta_n}+\widetilde{\sigma}^2_{\theta+(i-1)r_n\Delta_n}-1\big)_{i\in\mathcal{I}_2}\big)$. Then, the testing problem can be interpreted as a high dimensional location signal detection problem in the sup-norm. More precisely, we test the hypothesis

$$H_0:\sup_j(\tilde{\sigma}_j^2-1)=0\quad\text{against the alternative}\quad H_1:\sup_j(\tilde{\sigma}_j^2-1)\ge b_n\,,$$

and we are interested in the maximal value $b_n \to 0$ such that the hypothesis H_0 and H_1 are non-distinguishable in the minimax sense. Non-distinguishability in the minimax sense is formulated as

$$\lim_{n\to\infty} \inf_{\psi} \gamma_\psi\big(\mathfrak{a}, b_n\big) = 1, \tag{71}$$

and the detection boundary here is $b_n \propto (k_n\Delta_n)^{\mathfrak{a}} \propto n^{-\frac{\mathfrak{a}}{2\mathfrak{a}+1}}$. In order to show (71), we proceed in the fashion of Section 3.3.7 of Ingster and Suslina (2003). Let $\mathbb{P}_{\boldsymbol{\xi}}$ be the law of the observations. We consider the probability measures

$$\mathbb{P}_0 = \mathbb{P}_{\boldsymbol{\xi}} \times \mathbb{P}_{\theta_0} \quad \text{and} \quad \mathbb{P}_1 = \mathbb{P}_{\boldsymbol{\xi}} \times \mathbb{P}_{\theta_1},$$

where $\mathbb{P}_{\theta_0}$ means the hypothesis of the test applies (no jump) and $\mathbb{P}_{\theta_1}$ draws a jump-time θ with $\theta n k_n^{-1} \in \{1, \ldots, \lfloor n/k_n \rfloor - 1\}$ uniformly from this set. Therefore, $\mathbb{P}_0$ represents the probability measure without signal, and $\mathbb{P}_1$ the measure where a signal is present. It then follows that

$$\inf_{\psi} \gamma_\psi\big(\mathfrak{a}, b_n\big) \geq 1 - \frac{1}{2}\big\|\mathbb{P}_1 - \mathbb{P}_0\big\|_{TV} \geq 1 - \frac{1}{2}\big|\mathbb{E}_{P_0}\big[L_{0,1}^2 - 1\big]\big|^{1/2},$$

with $L_{0,1} = d\mathbb{P}_1/d\mathbb{P}_0$ the likelihood ratio of the measures $\mathbb{P}_1$ and $\mathbb{P}_0$. For the validity of (71), it thus suffices to establish

$$\mathbb{E}_{\mathbb{P}_0}\big[L_{0,1}^2\big] \to 1 \quad \text{as } n \to \infty. \tag{72}$$

To this end, for given θ we denote with $u_i^{\theta n k_n^{-1}} = \widetilde{\sigma}^4_{\theta+(i-1)r_n\Delta_n}$ for $i \in \mathcal{I}_2$, $v_i^{\theta n k_n^{-1}} = (u_i^{1/2} - 1)r_n^{1/2}$. We first perform some preliminary computations. Denote with $\varphi_Y(x)$ the density function of a Gaussian random variable Y, not necessarily standard normal, and for $a, b \in \{1, \ldots, \lfloor n/k_n \rfloor - 1\}$

$$I_{a,b}(x,y) := \left(\prod_{i\in\mathcal{I}_2} \frac{\varphi_{\tilde{\xi}_i(u_i^a)^{1/2}+v_i^a}(x_i)}{\varphi_{\tilde{\xi}_i}(x_i)}\right)\left(\prod_{i\in\mathcal{I}_2} \frac{\varphi_{\tilde{\xi}_i(u_i^b)^{1/2}+v_i^b}(y_i)}{\varphi_{\tilde{\xi}_i}(y_i)}\right).$$

Then, we have that $I_{a,b} := \int I_{a,b}(x,y) \prod_{i\in\mathcal{I}_2} \varphi_{\tilde{\xi}_i}(x_i)dx_i \prod_{i\in\mathcal{I}_2} \varphi_{\tilde{\xi}_i}(y_i)dy_i = 1$. Next, for $a \in \{1, \ldots, \lfloor n/k_n \rfloor - 1\}$, consider

$$II_a(x) := \prod_{i\in\mathcal{I}_2} \left(\frac{\varphi_{\tilde{\xi}_i(u_i^a)^{1/2}+v_i^a}(x_i)}{\varphi_{\tilde{\xi}_i}(x_i)}\right)^2.$$

Observe that for a standard Gaussian random variable Z and $s, t \in \mathbb{R}$, $|s| < 2$, we have

$$\mathbb{E}\big[\exp(sZ^2 + tZ)\big] = \frac{\exp\left(\frac{t^2}{2-4s}\right)}{\sqrt{1-2s}}. \tag{73}$$

This, together with the inequality

$$C_0 k_n \left(k_n\Delta_n\right)^{2\mathfrak{a}} \leq r_n \sum_{i=0}^{k_n/r_n-1} \left((k_n - ir_n)\Delta_n\right)^{2\mathfrak{a}} \leq k_n \left(k_n\Delta_n\right)^{2\mathfrak{a}}$$

for some constant $C_0 > 0$ and routine calculations yield that for some $C_0 \leq C_1 \leq 1$:

$$II_a := \int II_a(x) \prod_{i \in \mathcal{I}_2} \varphi_{\tilde{\xi}_i}(x_i) dx_i \leq e^{C_1 k_n (k_n \Delta_n)^{2\mathfrak{a}}} \big(1 + \mathcal{O}(1)\big).$$

With all the preliminary calculations completed, we are now ready to derive a bound for

$$\begin{aligned}\mathbb{E}_{\mathbb{P}_0}\big[L_{0,1}^2\big] - 1 = &\sum_{\substack{a,b=1\\ a\neq b}}^{\lfloor n/k_n \rfloor - 1} \mathbb{P}\big(\theta n k_n^{-1} = a\big)\mathbb{P}\big(\theta n k_n^{-1} = b\big)\big(I_{a,b} - 1\big) \\ &+ \sum_{a=1}^{\lfloor n/k_n \rfloor - 1} \mathbb{P}\big(\theta n k_n^{-1} = a\big)^2 \big(II_a - 1\big)\end{aligned}$$

where the first sum vanishes. For an appropriate choice of $c_k > 0$, $k_n = c_k\big(\sqrt{\log(m_n)}n^{\mathfrak{a}}\big)^{\frac{2}{2\mathfrak{a}+1}}$ is equivalent to $k_n(k_n\Delta_n)^{2\mathfrak{a}} = C_2 \log(n/k_n)$ for some $C_2 < C_1^{-1}$. Since $\mathbb{P}\big(\theta n k_n^{-1} = a\big) = k_n \Delta_n$, we thus obtain

$$\begin{aligned}\big|\mathbb{E}_{P_0}\big[L_{0,1}^2\big] - 1\big| &\leq \sum_{a=1}^{\lfloor n/k_n \rfloor} \mathbb{P}\big(\theta n k_n^{-1} = a\big)^2 \big(e^{C_1 k_n (k_n\Delta_n)^{2\mathfrak{a}}} - 1\big) \\ &= (1 + \mathcal{O}(1)) k_n \Delta_n \, e^{C_1 k_n (k_n \Delta_n)^{2\mathfrak{a}}}. \qquad (74)\end{aligned}$$

Using

$$k_n \Delta_n e^{C_1 k_n (k_n \Delta_n)^{2\mathfrak{a}}} = k_n \Delta_n \exp\big(C_1\, C_2 \log\big(n/k_n\big)\big) = (k_n \Delta_n)^{1 - C_1\, C_2} = \mathcal{O}(1),$$

we conclude (72).

Case $\mathfrak{a} \leq 1/2$: The only time we make use of the condition $\mathfrak{a} > 1/2$ above is in Step 3 to obtain $n/k_n^2 = \mathcal{O}(1)$. The necessity of this relation is due to the large number of blocks n/k_n, when operating with the entropy bounds. To establish the lower bound, this constraint can be removed by granting the experimenter even more additional information what is briefly sketched in the following. Indeed, suppose we know in addition that $\theta n \in \big\{k_n, 2k_n, \ldots, l_n k_n\big\}$ where $l_n = n^{\mathfrak{l}} \ll n/k_n$, $\mathfrak{l} > 0$ arbitrarily small but strictly positive and such that $l_n \in \mathbb{N}$. Using the sufficiency argument of Step 2, we can gather all the information contained in $(\eta_i)_{l_n k_n < i \leq n}$ in one single average $(n - (l_n + 1)k_n)^{-1} \sum_{i = l_n k_n + 1}^{n} \eta_i^2$. Then, one can repeat Steps 3 and 4, subject to the weaker condition $l_n/k_n = \mathcal{O}(1)$. Selecting $\mathfrak{l} > 0$ sufficiently small for each $0 < \mathfrak{a} \leq 1$, this is always possible. The lower bound in Step 4 gives the same minimax detection boundary and hence the claim follows.

Let us now touch on the general case with some $\sigma_-^2 > 0$ and sequences L_n. Exactly the same arguments lead to (with c_k as before) $\lim_{n\to\infty} \inf_\psi \gamma_\psi(\mathfrak{a}, b_n) = 1$ for

$$b_n \leq L_n\, (k_n \Delta_n)^{\mathfrak{a}}, \quad k_n = c_k\Big(\sqrt{\log(m_n)}\sigma_-^2 L_n^{-1} n^{\mathfrak{a}}\Big)^{\frac{2}{2\mathfrak{a}+1}} = c_k\, n^{\frac{2\mathfrak{a}}{2\mathfrak{a}+1}} \left(\frac{\sigma_-^4 \log(m_n)}{L_n^2}\right)^{\frac{1}{2\mathfrak{a}+1}}, \quad (75)$$

which gives the general result.

Finally, we remark on the regularity alternatives $H_1^{\mathrm{R}1/2}$. The strategy for the proof is exactly the same for both cases and along the same lines as for jumps above. Instead of a jump of size $L_n(k_n\Delta_n)^{\mathfrak{a}}$

at unknown location, we observe a sudden, more regular increase in σ_t^2, where we exploit the regularity $\mathfrak{a}'$. Hence, the jump gets replaced with a gradual regular increase. However, since $\mathfrak{a}' < \mathfrak{a}$, this is always possible on an interval of size $(k_n\Delta_n)^{\mathfrak{a}/\mathfrak{a}'} < k_n\Delta_n$, and thus the arguments are almost identical. This also highlights the fact that at (or below) the boundary b_n, all three alternatives H_1^{J} and $H_1^{\mathrm{R}1/2}$ are not distinguishable. □

Proof of Theorem 4.3. Using similar arguments as in the proof of Theorem 3.2 and in particular Proposition A.4 one derives for

$$\overline{V}_{n,i} = \left| \frac{\frac{n}{k_n^\diamond}\sum_{j=i-k_n^\diamond+1}^{i}\big((\Delta_j^n X)^2 - \mathbb{E}\big[(\Delta_j^n X)^2\big]\big)}{\frac{n}{k_n^\diamond}\sum_{j=i+1}^{i+k_n^\diamond}\big((\Delta_j^n X)^2 - \mathbb{E}\big[(\Delta_j^n X)^2\big]\big)} - 1 \right|, \quad k_n^\diamond \le i \le n - k_n^\diamond\,,$$

that under the alternatives H_1^{J} and $H_1^{\mathrm{R}1/2}$, it holds that

$$\sqrt{k_n^\diamond}\,\overline{V}_{n,i} = \mathcal{O}_P(1), \quad k_n^\diamond \le i \le n - k_n^\diamond. \tag{76}$$

Based on (76), a simple estimate yields

$$\begin{aligned} V_n^* &\ge -\overline{V}_{n,\lfloor n\theta\rfloor} + \frac{n}{k_n^\diamond}\left|\int_{\theta-k_n^\diamond\Delta_n}^{\theta}\sigma_s^2\,ds - \int_\theta^{\theta+k_n^\diamond\Delta_n}\sigma_s^2\,ds\right|\frac{\big(1-\mathcal{O}_{\mathbb{P}}(1)\big)}{\sigma_\theta^2} \\ &\ge -\mathcal{O}_{\mathbb{P}}\big((k_n^\diamond)^{-1/2}\big) + \frac{n}{k_n^\diamond}\left|\int_{\theta-k_n^\diamond\Delta_n}^{\theta}\sigma_s^2\,ds - \int_\theta^{\theta+k_n^\diamond\Delta_n}\sigma_s^2\,ds\right|\frac{\big(1-\mathcal{O}_{\mathbb{P}}(1)\big)}{\sup_{0\le t\le 1}\sigma_t^2}. \end{aligned} \tag{77}$$

Observe that in order to prove $\gamma_{\psi^\diamond}(\mathfrak{a}, b_n^\diamond) \to 0$, it suffices to show that

$$\mathbb{P}\Big(V_n^* \ge 2C^\diamond\sqrt{2\log(m_n^\diamond)/k_n^\diamond}\Big) \to 1 \quad \text{under } H_1^{\mathrm{J}} \text{ or } H_1^{\mathrm{R}1/2}, \tag{78}$$

$$\text{and } \mathbb{P}\Big(V_n^* < 2C^\diamond\sqrt{2\log(m_n^\diamond)/k_n^\diamond}\Big) \to 1 \quad \text{under } H_0. \tag{79}$$

Case H_1^{J}: Using that $(\sigma_t^2 - \Delta\sigma_t^2)_{t\in[0,1]} \in \Sigma(\mathfrak{a}, L_n)$ and $\Delta\sigma_t \ge 0$, we get

$$\frac{n}{k_n^\diamond}\sup_{t\ge\theta}\left|\int_{\theta-k_n^\diamond\Delta_n}^{\theta}\sigma_s^2\,ds - \int_\theta^{\theta+k_n^\diamond\Delta_n}\sigma_s^2\,ds\right| \ge b_n - 2L_n\,(k_n^\diamond\Delta_n)^{\mathfrak{a}}\,.$$

Hence (78) follows for (34) with (77).

Case $H_1^{\mathrm{R}1/2}$: We consider $H_1^{\mathrm{R}1}$ here, the proof for $H_1^{\mathrm{R}2}$ follows the same strategy and is omitted. For $\sigma_t^2 \in \mathcal{S}_\theta^{\mathrm{R}1}\big(\mathfrak{a}, \mathfrak{a}', b_n^\diamond, L_n, C\big)$, we have that

$$\sigma_{\theta+h}^2 \ge \sigma_\theta^2 + Ch^{\mathfrak{a}'} \quad \text{or} \quad \sigma_{\theta+h}^2 \le \sigma_\theta^2 - Ch^{\mathfrak{a}'}, \quad \text{for } 0 \le h \le (b_n^\diamond)^{1/\mathfrak{a}}.$$

Since $(b_n^\diamond)^{1/\mathfrak{a}} \ge k_n^\diamond\Delta_n$ and $(\sigma_{t\wedge\theta}^2)_{t\in[0,1]} \in \Sigma(\mathfrak{a}, L_n)$, it follows that

$$\begin{aligned} \frac{n}{k_n^\diamond}\left|\int_{\theta-k_n^\diamond\Delta_n}^{\theta}\sigma_s^2\,ds - \int_\theta^{\theta+k_n^\diamond\Delta_n}\sigma_s^2\,ds\right| &\ge \frac{C\,n}{k_n^\diamond}\int_0^{k_n^\diamond\Delta_n} s^{\mathfrak{a}'}\,ds - L_n\,(k_n^\diamond\Delta_n)^{\mathfrak{a}} \\ &\ge \frac{C}{1+\mathfrak{a}'}\,(k_n^\diamond\Delta_n)^{\mathfrak{a}'} - L_n\,(k_n^\diamond\Delta_n)^{\mathfrak{a}}\,. \end{aligned}$$

Thus, using $L_n = \mathcal{O}\big((n/k_n^\diamond)^{\mathfrak{a}-\mathfrak{a}'}\big)$, for sufficiently large n the above is bounded from below by $b_n^\diamond$. Therefore (78) follows with (77).

Case H_0: Under the hypothesis we employ the upper bound

$$V_n^* \le \max_{k_n^\diamond \le i \le n-k_n^\diamond} \overline{V}_{n,i} + L_n(k_n^\diamond \Delta_n)^{\mathfrak{a}}$$

to prove (79). (33) implies

$$2C^\diamond\sqrt{2\log(m_n^\diamond)/k_n^\diamond} \ge 2\sqrt{2\log(m_n^\diamond)/k_n^\diamond} + L_n(k_n^\diamond\Delta_n)^{\mathfrak{a}},$$

and hence it suffices to show that $\mathbb{P}\big(\max_{k_n^\diamond \le i \le n-k_n^\diamond} \overline{V}_{n,i} \le \sqrt{2\log(m_n^\diamond)/k_n^\diamond}\big) \to 1$. This, however, follows from a very easy adaption of Theorem 3.2. Hence (79) follows, which completes the proof. □

Appendix D: Proofs of Section 5

Proof of Proposition 5.1. We use the following elementary lemma to prove Proposition 5.1.

Lemma D.1. *Let $f(t)$ and $g(t)$ be functions on $[0,\theta]$ such that $f(t)$ is increasing. As long as $f(\theta) - f(\theta-\gamma) \ge \sup_{0\le t\le\theta}|g(t)|$ for some $\gamma \in [0,\theta]$, we have that*

$$\operatorname{argmax}_{0\le t\le\theta}\big(f(t)+g(t)\big) \ge \theta - \gamma.$$

An analogous result holds if $f(t)$ and $g(t)$ are functions on $[\theta,1]$ and $f(t)$ is decreasing.

For $\theta \in (0,1)$ define $i^* = \lceil \theta n \rceil$, the smallest integer such that $i^*\Delta_n$ is larger or equal than θ. While $(\sigma_t^2)_{t\in[0,1]}$ is the squared volatility process containing one jump at time θ, denote by $(\tilde\sigma_t^2)_{t\in[0,1]}$ the same path without jump, such that

$$\sigma_{i\Delta_n}^2 = \tilde\sigma_{i\Delta_n}^2 + \delta \mathbb{1}(i \ge i^*)$$

with jump size δ. Without loss of generality, we assume $\delta > 0$. Define

$$f\big(i\Delta_n\big) = \begin{cases} 0, & \text{if } i + k_n < i^*, \\ (i+k_n-i^*)k_n^{-1/2}\delta & \text{for } i = i^*-k_n,\dots,i^*\,, \\ \sqrt{k_n}\delta & \text{if } i > i^*, \end{cases} \tag{80}$$

and $(f(t))_{t\in[0,1]}$ the associated piecewise constant increasing step function. For $i = k_n,\dots,n-k_n$:

$$\begin{aligned}
&\sum_{j=i-k_n+1}^{i} n(\Delta_j^n X)^2 - \sum_{j=i+1}^{i+k_n} n(\Delta_j^n X)^2 \\
&= \Big\{\sum_{j=i-k_n+1}^{i}\big(n(\Delta_j^n X)^2 - \mathbb{E}[n(\Delta_j^n X)^2]\big) - \sum_{j=i+1}^{i+k_n}\big(n(\Delta_j^n X)^2 - \mathbb{E}[n(\Delta_j^n X)^2]\big)\Big\} \\
&+ \Big\{\sum_{j=i-k_n+1}^{i}\big(\mathbb{E}[n(\Delta_j^n X)^2] - \tilde\sigma_{j\Delta_n}^2\big) - \sum_{j=i+1}^{i+k_n}\big(\mathbb{E}[n(\Delta_j^n X)^2 - \sigma_{j\Delta_n}^2]\big)\Big\} \\
&+ \Big\{\sum_{j=i-k_n+1}^{i}\tilde\sigma_{j\Delta_n}^2 - \sum_{j=i+1}^{i+k_n}\tilde\sigma_{j\Delta_n}^2\Big\} - \sum_{j=i+1}^{i+k_n}\big(\sigma_{j\Delta_n}^2 - \tilde\sigma_{j\Delta_n}^2\big) =: A_i^n + B_i^n + C_i^n - \sum_{j=i+1}^{i+k_n}\big(\sigma_{j\Delta_n}^2 - \tilde\sigma_{j\Delta_n}^2\big),
\end{aligned}$$

with the obvious definition using the curly brackets. Thus, for the step function $(g(t))_{t\in[0,1]}$, with

$$g\big(i\Delta_n\big)=k_n^{-1/2}\Bigg(\sum_{j=i-k_n+1}^{i}n(\Delta_j^nX)^2-\sum_{j=i+1}^{i+k_n}n(\Delta_j^nX)^2+\sum_{j=i+1}^{i+k_n}\big(\sigma^2_{i^*\Delta_n}-\tilde\sigma^2_{i^*\Delta_n}\big)\Bigg)$$

for $i=k_n,\dots,n-k_n$ and $g(i\Delta_n)=0$ else, we have that

$$\begin{aligned}\sqrt{k_n}\,g\big(i\Delta_n\big)&=A_{n,i}+B_{n,i}+C_{n,i}+D_{n,i}\,,\\ -D_{n,i}&=\sum_{j=i+1}^{i+k_n}\big(\sigma^2_{j\Delta_n}-\tilde\sigma^2_{j\Delta_n}\big)-\big(\sigma^2_{i^*\Delta_n}-\tilde\sigma^2_{i^*\Delta_n}\big)\\ &=\sum_{j=i+1}^{i+k_n}\big(\sigma^2_{j\Delta_n}-\sigma^2_{i^*\Delta_n}\big)+\big(\tilde\sigma^2_{i^*\Delta_n}-\tilde\sigma^2_{j\Delta_n}\big)\,.\end{aligned}$$

Exploiting the smoothness of $(\sigma_t)_{t\in[0,1]}$ by Assumption 3.1, we obtain

$$\begin{aligned}\max_{i=1,\dots,n}\big|D_{n,i}\big|&=\max_{i^*-k_n\le i\le i^*}\big|D_{n,i}\big|\le K\sup_{0\le s\le 1}|\sigma_s|\max_{i^*-k_n\le i\le i^*}\sum_{j=i+1}^{i+k_n}|\sigma_{j\Delta_n}-\sigma_{i^*\Delta_n}|\\ &\le K\sup_{0\le s\le 1}|\sigma_s|k_n^{1+\mathfrak{a}}n^{-\mathfrak{a}}=\mathcal{O}_{a.s.}\big(\sqrt{k_n\log(n)}\big)\,,\end{aligned}$$

with some constant K by (17). Proceeding similarly as in the proof of Theorem 3.2, it follows that

$$\max_{i=1,\dots,k_n}\big|A_{n,i}+B_{n,i}+C_{n,i}\big|=\mathcal{O}_{\mathbb{P}}\big(\sqrt{k_n\log(n)}\big)\,. \tag{81}$$

Altogether, we conclude that

$$\sup_{t\in[0,\theta]}\big|g(t)\big|=\mathcal{O}_{\mathbb{P}}\big(\sqrt{\log(n)}\big)\,. \tag{82}$$

Finally, using (80), we see that $f(i\Delta_n)>|g(i\Delta_n)|>0$ holds for each $i=i^*-k_n/2,\dots,i^*$, with probability tending to 1. In particular,

$$V^{\diamond}_{n,i}=\big|f\big(i\Delta_n\big)+g\big(i\Delta_n\big)\big|=f\big(i\Delta_n\big)+\operatorname{sign}\big(g(i\Delta_n)\big)\big|g\big(i\Delta_n\big)\big| \tag{83}$$

for those i. Furthermore,

$$f(i^*\Delta_n)-f(i^*\Delta_n-\gamma_n)=\lfloor\gamma_n n\rfloor\delta k_n^{-1/2}\ \text{ for }\ \gamma_n\in[0,k_n/(2n)]\,.$$

Thus, we choose γ_n such that

$$\frac{\sqrt{k_n\log(n)}}{\delta n}=\mathcal{o}\big(\gamma_n\big). \tag{84}$$

Now the assumptions of Lemma D.1 are fulfilled, and we obtain, with probability tending to one,

$$i^*\Delta_n\ge\operatorname{argmax}_{i=k_n,\dots,i^*}V^{\diamond}_{n,i}\,\Delta_n\ge i^*\Delta_n-\gamma_n,$$

through an application of Lemma D.1 together with (83). A similar argument for $i > i^*$ shows

$$i^*\Delta_n \leq \text{argmax}_{i=i^*,\dots,n-k_n} V^{\diamond}_{n,i}\,\Delta_n \leq i^*\Delta_n + \gamma_n,$$

from which one obtains $|\widehat{\theta}_n - i^*\Delta_n| = \mathcal{O}_{\mathbb{P}}\big(\gamma_n\big)$, which completes the proof by definition of i^*. □

Proof of Proposition 5.3. Suppose that

$$\sqrt{k_n}\delta_n = o\big(\sqrt{\log(n)}\big) \tag{85}$$

and we have a consistent estimator $\widehat{\theta}^*$ for θ. Define

$$T_{\widehat{\theta}^*} = \frac{n}{k_n}\Bigg(\sum_{j=\widehat{\theta}^* n-k_n}^{\widehat{\theta}^* n-1}\big(\Delta_j^n X\big)^2 - \sum_{j=\widehat{\theta}^* n+1}^{\widehat{\theta}^* n+k_n}\big(\Delta_j^n X\big)^2\Bigg). \tag{86}$$

Using the statistic $T_{\widehat{\theta}^*}$, we can now test for jumps in the volatility σ_t^2. Note that due to (85), it readily follows that this new test has a detection boundary $b_n = o\big(\sqrt{\log(n)}/\sqrt{k_n}\big)$. This, however, is a contradiction to Theorem 4.1, and hence such an estimator $\widehat{\theta}^*$ cannot exist. □

References

Aït-Sahalia, Y. and J. Jacod (2009). Testing for jumps in a discretely observed process. *The Annals of Statistics 37*(1), 184–222.

Aït-Sahalia, Y. and J. Jacod (2010). Is Brownian motion necessary to model high-frequency data? *The Annals of Statistics 38*(5), 3093–3128.

Alvarez, A., F. Panloup, M. Pontier, and N. Savy (2012). Estimation of the instantaneous volatility. *Statistical Inference for Stochastic Processes 15*(1), 27–59.

Andersen, T. G. and T. Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review 39*, 885–905.

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica 61*(4), 821–56.

Aue, A., S. Hörmann, L. Horváth, and M. Reimherr (2009). Break detection in the covariance structure of multivariate time series models. *The Annals of Statistics 37*(6B), 4046–4087.

Bai, J. and P. Perron (1998). Estimating and testing linear models with multiple structural changes. *Econometrica 66*(1), 47–78.

Barndorff-Nielsen, O. E. and N. Shephard (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society 64*(2), 253–280.

Bobkov, S. G., G. P. Chistyakov, and F. Götze (2013). Rate of convergence and Edgeworth-type expansion in the entropic central limit theorem. *The Annals of Probability 41*(4), 2479–2512.

Brown, L. D. and M. G. Low (1996). Asymptotic equivalence of nonparametric regression and white noise. *The Annals of Statistics 24*(6), 2384–2398.

Csörgő, M. and L. Horváth (1997). *Limit theorems in change-point analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester. With a foreword by David Kendall.

Gatheral, J., T. Jaisson, and M. Rosenbaum (2014). Volatility is rough. *Preprint, arXive: 1410.3394*.

Hinkley, D. V. (1971). Inference about the change-point from cumulative sum tests. *Biometrika 58*, 509–523.

Hoffmann, M. and R. Nickl (2011). On adaptive inference and confidence bands. *The Annals of Statistics 39*(5), 2383–2409.

Iacus, S. M. and N. Yoshida (2012). Estimation for the change point of volatility in a stochastic differential equation. *Stochastic Processes and their Applications 122*(3), 1068–1092.

Ingster, Y. I. (1993). Asymptotically minimax hypothesis testing for nonparametric alternatives i, ii, iii. *Mathematical Methods of Statistics 2*(4), 85–114; 171–189; 249–268.

Ingster, Y. I. and I. A. Suslina (2003). *Nonparametric goodness-of-fit testing under Gaussian models*, Volume 169 of *Lecture Notes in Statistics*. Springer-Verlag, New York.

Jacod, J. (1997). On continuous conditional Gaussian martingales and stable convergence in law. *Séminaire de Probabilitiés, Strasbourg, tome 31*, 232–246.

Jacod, J. (2008). Asymptotic properties of realized power variations and related functionals of semimartingales. *Stochastic Processes and their Applications 118*(4), 517–559.

Jacod, J. and P. Protter (2012). *Discretization of processes*. Springer.

Jacod, J. and M. Rosenbaum (2013). Quarticity and other functionals of volatility: efficient estimation. *The Annals of Statistics 41*(3), 1462–1484.

Jacod, J. and V. Todorov (2010). Do price and volatility jump together? *The Annals of Applied Probability 20*(4), 1425–1469.

Komlós, J., P. Major, and G. Tusnády (1975). An approximation of partial sums of independent RV's and the sample DF. I. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 32*, 111–131.

Komlós, J., P. Major, and G. Tusnády (1976). An approximation of partial sums of independent RV's, and the sample DF. II. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 34*(1), 33–58.

Loader, C. R. (1996). Change point estimation using nonparametric regression. *The Annals of Statistics 24*(4), 1667–1678.

Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics 36*(4), 270–296.

Marsaglia, G., W. W. Tsang, and J. Wang (2003). Evaluating Kolmogorov's distribution. *Journal of Statistical Software 8*(18), 1–4.

Müller, H.-G. (1992). Change-points in nonparametric regression analysis. *The Annals of Statistics 20*(2), 737–761.

Müller, H.-G. and U. Stadtmüller (1999). Discontinuous versus smooth regression. *The Annals of Statistics 27*(1), 299–337.

Mykland, P. and L. Zhang (2009). Inference for continuous semimartingales observed at high frequency. *Econometrica 77*(5), 1403–1445.

Mykland, P. A. (2012). A Gaussian calculus for inference from high frequency data. *Annals of Finance 8*, 235–258.

Page, E. S. (1955). A test for a change in a parameter occurring at an unknown point. *Biometrika 42*, 523–527.

Pettitt, A. N. (1980). A simple cumulative sum type statistic for the change-point problem with zero-one observations. *Biometrika 67*(1), 79–84.

Phillips, P. C. B. (1987). Time series regression with a unit root. *Econometrica 55*(2), 277–301.

Sakhanenko, A. I. (1996). Estimates for the accuracy of constructions on a probability space in the central limit theorem. *Rossiĭskaya Akademiya Nauk. Sibirskoe Otdelenie. Institut Matematiki im. S. L. Soboleva. Sibirskiĭ Matematicheskiĭ Zhurnal 37*(4), 919–931, iv.

Spokoiny, V. G. (1998). Estimation of a function with discontinuities via local polynomial fit with an adaptive window choice. *The Annals of Statistics 26*(4), 1356–1378.

Tauchen, G. and V. Todorov (2011). Volatility jumps. *Journal of Business and Economic Statistics 29*, 356–371.

Vetter, M. (2012). Estimation of correlation for continuous semimartingales. *Scandinavian Journal of Statistics 39*(4), 757–771.

Wu, W. B. (2007). Strong invariance principles for dependent random variables. *The Annals of Probability 35*(6), 2294–2320.

Wu, W. B. and Z. Zhao (2007). Inference of trends in time series. *Journal of the Royal Statistical Society. Series B. Statistical Methodology 69*(3), 391–410.

Zaĭtsev, A. Y. (1987). On the Gaussian approximation of convolutions under multidimensional analogues of S. N. Bernstein's inequality conditions. *Probability Theory and Related Fields 74*(4), 535–566.

SFB 649 Discussion Paper Series 2015

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Pricing Kernel Modeling" by Denis Belomestny, Shujie Ma and Wolfgang Karl Härdle, January 2015.

002 "Estimating the Value of Urban Green Space: A hedonic Pricing Analysis of the Housing Market in Cologne, Germany" by Jens Kolbe and Henry Wüstemann, January 2015.

003 "Identifying Berlin's land value map using Adaptive Weights Smoothing" by Jens Kolbe, Rainer Schulz, Martin Wersing and Axel Werwatz, January 2015.

004 "Efficiency of Wind Power Production and its Determinants" by Simone Pieralli, Matthias Ritter and Martin Odening, January 2015.

005 "Distillation of News Flow into Analysis of Stock Reactions" by Junni L. Zhang, Wolfgang K. Härdle, Cathy Y. Chen and Elisabeth Bommes, January 2015.

006 "Cognitive Bubbles" by Ciril Bosch-Rosay, Thomas Meissnerz and Antoni Bosch-Domènech, February 2015.

007 "Stochastic Population Analysis: A Functional Data Approach" by Lei Fang and Wolfgang K. Härdle, February 2015.

008 "Nonparametric change-point analysis of volatility" by Markus Bibinger, Moritz Jirak and Mathias Vetter, February 2015.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".